

The College at Brockport: State University of New York Digital Commons @Brockport

Education and Human Development Master's
Theses

Education and Human Development

Fall 12-4-2013

The Influence of the Reading First Grant Program on English Language Arts Classrooms

John F. Hoffman

The College at Brockport, johnhoff42@gmail.com

Follow this and additional works at: http://digitalcommons.brockport.edu/ehd_theses



Part of the [Education Commons](#)

To learn more about our programs visit: <http://www.brockport.edu/ehd/>

Repository Citation

Hoffman, John F., "The Influence of the Reading First Grant Program on English Language Arts Classrooms" (2013). *Education and Human Development Master's Theses*. 293.

http://digitalcommons.brockport.edu/ehd_theses/293

This Thesis is brought to you for free and open access by the Education and Human Development at Digital Commons @Brockport. It has been accepted for inclusion in Education and Human Development Master's Theses by an authorized administrator of Digital Commons @Brockport. For more information, please contact kmyers@brockport.edu.

The Influence of the Reading First Grant Program on English Language Arts Classrooms

By

John Fredrick Hoffman

December 30th, 2013

A project submitted to the Department of Education and Human Development of the State
University of New York College at Brockport in partial fulfillment of the requirements for the
degree of Master of Science in Education

The Influence of the Reading First Grant Program on English Language Arts Classrooms

By

John Fredrick Hoffman

APPROVED BY:

Advisor

Graduate Director

Date

Date

Acknowledgements

This thesis owes a large debt to the patient support of Brittany Taylor, the love of my life. Special thanks goes to my advisor, Dr. Thomas Giblin, whose wisdom and candid approach kept me honest, and to my parents, who patiently reviewed this work for accessibility.

Table of Contents

| | |
|--|----|
| Chapter 1: Problem Statement | 4 |
| Chapter 2: Literature Review | 12 |
| Chapter 3: Application | 48 |
| Chapter 4: Conclusions and Recommendations | 66 |
| References | 70 |
| Appendix A | 85 |
| Appendix B | 86 |
| Appendix C | 87 |
| Appendix D | 88 |

Chapter 1: Introduction

The federal Reading First (RF) grant program was established as part of the No Child Left Behind Act (NCLB). NCLB aimed to create systematic school improvement by providing some administrative flexibility and funding to state and local education agencies (SEAs and LEAs) in return for the requirements that they: create standards; assess progress toward those standards using standardized tests; and deploy institutional consequences for failure to meet standards. RF was the largest grant program created under NCLB, authorizing \$5.39 billion in grant funding between 2002 and 2008 (USDOE, 2009), and was intended to improve reading instruction in high-poverty K-3 LEAs by providing additional funding for professional development and curriculum materials that promoted “scientifically based instructional strategies” (“NCLB,” 2002, Sec. 1001 (9)).

Problem Statement

NCLB constituted unprecedented federal influence on education reform efforts (*No child left behind?*, 2003), and RF shaped the reform efforts that were to improve reading instruction for underprivileged and low-performing students in grades K-3 across the nation (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Scott, 2006). Strict grant requirements were enforced by federal and district officials alike, with select programs and techniques mandated to the exclusion of others and restriction of local adaptations. *Education Week* reporting (Manzo, 2007a) and government investigations (Ashby & General Accounting Office, 2007; USDOE, Office of Inspector General, 2006) both recognized USDOE mismanagement of the program. RF schools saw gains on state tests (Bean, Draper, Turner, & Zigmond, 2010), but RF schools did not demonstrate significant improvements over Title I schools in government evaluations (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). As a result, appropriations for RF were reduced

61% by Congress in fiscal year 2008 (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009). No scholarly work to date has analyzed the translation of RF from federal policy to classroom practice with regard to its impact on the English/language arts (ELA) curriculum. This synthesis of research analyzes the mandates, implementation, and outcomes of RF from 2002 to the program's dramatic attenuation in 2008 and closure in 2009 in order to assess its impact on the ELA curriculum.

Significance of the Problem

RF made significant changes to the curriculum among participating schools. Sixty-seven percent of districts participating in RF reported making changes to reading instruction (Scott, 2007b). Among these changes was an increase in instructional time to and beyond the 90 minute instructional block mandated by RF (Moss et al., 2008), with one study of reading instructional time finding an average daily allotment of 105.7 minutes of reading instruction in RF schools compared to 87.2 minutes in similar Title I schools (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). Classroom observations have found this resulted in a moderate, positive, and statistically significant increases in time observers identified as addressing the key components of reading instruction highlighted by RF (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008a). Under RF requirements, schools widely adopted prepackaged reading programs that script reading lessons (Dudley-Marling, 2005), including more highly explicit instruction (B. C. Gamse, Jacob, Horst, Boulay, & Unlu, 2008), and a legally-tenuous emphasis on small group work (Healy, 2007; USDOE, Office of Inspector General, 2006).

About \$1 billion was appropriated for RF each year until 2007, creating a total investment of more than \$5.39 billion in improving reading instruction (USDOE, 2009), or \$743 annually per child served (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich,

2009). The funds provided directly affected 1,809 districts and 5,880 schools near the program's peak in April 2007 (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b), influencing schools in all 50 states, the District of Columbia, Bureau of Indian Affairs schools, and those in American Samoa, the Commonwealth of Puerto Rico and U.S. Virgin Islands (Callow-Heusser & Chapman, 2007; McCallion, 2008; Scott, 2007b; USDOE, Office of Elementary and Secondary Education, 2002). This constitutes 7-8% of districts nationwide and around 4% of elementary schools according to conservative estimates (Ashby & General Accounting Office, 2007; Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009), but half of participating districts reported implementing features of RF in non-RF schools or higher grades (Scott, 2007b). Scott's (2007b) survey of states suggests that at least 3,274 non-RF districts participated in RF professional development provided by states, and 80% of states and 75% of districts with RF reported coordination of Title I and RF requirements, which both may have served to extend the reach of RF features to further districts.

The pressures NCLB applied to schools to improve (i.e., by requiring purchase of additional instructional services, and threatening staff with termination, institutional restructuring, and school closure) may have motivated schools to adopt solutions portrayed by USDOE as "scientifically based reading instruction" (SBRR; USDOE, Office of Elementary and Secondary Education, 2002). This incentive to improve student achievement scores provided by AYP requirements and legitimation of RF program methods as SBRR by USDOE may have prompted non-RF schools to adopt RF core reading programs and curriculum features. The relative ease of implementing the scripted programs supported under RF compared to developing district-specific programs likely enhanced the appeal of RF feature adoption. The coordination of

RF and Title I funding among some districts may have served as a means to disseminate RF program features (Scott & Fagan, 2005).

The influence of RF therefore contributed to increased time allotted for reading instruction, more direct instruction through scripted reading programs, and small group work in the form mandated by the legislation. Analysis of RF's mandates, implementation, and outcomes of the RF program in relation to the ELA curriculum will allow educators and policymakers to understand instructional shifts brought about by RF programs and adjust instruction and policy.

Purpose

The purpose of this work is to analyze the mandates, implementation, and outcomes of RF to assess the quality and extent of changes in the ELA curriculum. This analysis will provide insight into the impact of RF on the ELA curriculum, the outcomes of RF programs, and the process of the reform as it was translated from federal policy to classroom practice.

Rationale

This study will inform teachers of curriculum shifts that resulted from RF and their impacts; provide insight for policymakers in the ways the RF mandates translated into classroom practice; assess RF's impact on curriculum, teachers, and students so that further efforts at school reform can be revised, and provide suggestions for further study.

Roadmap

Chapter 1 of this volume provides an introduction to this work, a statement of the problem it addresses, the significance of the problem, and a purpose and rationale for addressing the problem. Chapter 2 will present a literature review of the mandates, implementation, and outcomes of RF policy. Chapter 3 will present an analysis of the effect and quality of RF's influence on the ELA curriculum and how it translated from federal policy to classroom practice.

Chapter 4 will present conclusions reached about the influence of RF, its effects, and recommendations for future research.

Definition of Terms

Reading First (RF)—A United States Federal Government grant program created under the No Child Left Behind Act to fund “scientifically based” reading instructional programs serving grades K-3 in high-poverty schools. RF was intended to model effective instruction for states and districts that failed to meet “adequate yearly progress” under No Child Left Behind.

No Child Left Behind Act (NCLB)—The 2002 reauthorization of the Elementary and Secondary Education Act of 1965. NCLB required states to set content and performance standards for ELA and Math, and employ standardized testing to assess students’ progress towards those standards. States were required to set improvement goals, known as “adequate yearly progress,” sufficient to achieve 100% proficiency on standardized tests by 2014. Institutional and individual consequences were attached to failure to meet these goals, including public school choice, firing key district personnel, school restructuring, and state takeover.

Scientifically Based Reading Research (SBRR)—Research designated by USDOE as “rigorous, systematic, and objective” study of reading instruction that “employs systematic, empirical methods that draw on observation or experiment” to “obtain valid knowledge relevant to reading development, reading instruction, and reading difficulties” under NCLB [USDOE, 2002, sec. 9101 (37)].

United States Department of Education (USDOE)—The federal agency tasked with overseeing national education policy. USDOE was responsible for the process and approval of states’ RF applications; contracting for and providing technical assistance; monitoring state and

local implementation of RF; and contracting for and conducting evaluation studies (USDOE, 2002).

National Reading Panel (NRP)—A body commissioned by the United States Congress in 1997 to assess the effectiveness of different methods of teaching reading by reviewing reading research (National Institute of Child Health and Human Development, 2000).

State Educational Agency (SEA)—State level government agencies tasked with oversight and support of schools and residents in educational matters.

Local Educational Agency (LEA)—An organization that operates local public primary and secondary schools.

Adequate Yearly Progress (AYP)—The goal improvement level set by states under NCLB so that schools reach 100% proficiency on state performance standards by 2014. Success at reaching these levels of improvement determined whether consequences including firings, school restructurings, or state takeovers occurred under NCLB (*No child left behind?*, 2003)

Title I of the Elementary and Secondary Education Act (Title I)—A federal program that provides LEAs financial support targeted to areas with high concentrations of poverty to help provide adequate education (USDOE, 2011)

Dynamic Indicators of Basic Early Literacy Skills (DIBELS)—A series of tests times for one minute each developed by scholars at the University of Oregon to evaluate students literacy skills. Subtests are designed to evaluate phonics skills, fluency, vocabulary, and comprehension (Li & Zhang, 2008)

Response to Intervention (RTI)—A three-tiered system for addressing students' poor performance in academic tasks. Frequent assessments indicate whether students may be falling

behind and greater levels of support are provided for students until they improve (National Center on Response to Intervention, n.d.).

The Observation Survey of Early Literacy Achievement (OSEL A)—OSEL A is an assessment of literacy skills that is informal and designed to help teachers interpret student needs as part of the Reading Recovery program. The test is untimed, featuring a running record and evaluations of students' knowledge of print conventions, reading skills, vocabulary and phonics skills (Li & Zhang, 2008).

Group Reading Assessment and Diagnostic Evaluation (GRA+DE)—Assesses what developmental skills students display (Pearson Education, 2013a) o

Stanford Achievement Test Ninth Edition/Tenth Edition (SAT-9/SAT-10)—A test popular in the United States as a requirement for college admission, SAT includes measures of reading, writing, math, and science skills.

Peabody Picture Vocabulary Test Second Edition (PPVT-II)—A test of vocabulary and word retrieval (Pearson Education, 2013b).

Iowa Test of Basic Skills (ITBS)—A series of test that measure students' skills in specific content areas. Throughout this work, ITBS will refer to the reading assessments employed to evaluate RF programs (Houghton Mifflin Harcourt, 2010).

TerraNova—A series of assessments covering reading, language arts, math, science, and social studies. Throughout this work, TerraNova will refer to the reading assessments employed by some states to evaluate RF performance (CTB/McGraw-Hill LLC., 2013)

National Assessment of Educational Progress (NAEP)—A nationally representative assessment of students abilities including tests of mathematics, reading, writing science, the arts, civics, economics, geography, and U.S. history. NAEP does not provide individual scores, but

aggregates data that allows inferences about states, subgroups, and national trends (National Center for Education Statistics, 2012)

Chapter 2: Literature Review

Introduction

The RF Program. RF was designed to help high-poverty and low-literacy schools improve reading instruction in grades K-3 by providing funding for professional development and curriculum materials that promoted instruction based on scientific research (USDOE, 2009). RF provided discretionary grants to states, who retained up to 20% of funds for state-level professional development and grant administration, and provided the remainder to districts through a competitive sub-grant process¹ (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009; Robelen, 2002; Scott & Fagan, 2005). Grant requirements were structured by the National Reading Panel (National Institute of Child Health and Human Development, 2000)'s framework of reading instruction and required certain curricular features and reporting at the state and local levels (Cunningham, 2001; Healy, 2007; Morrow, Rueda, & Lapp, 2009).

NCLB Context. President George W. Bush the NCLB Act² into law in January 2002, with bipartisan support in reauthorizing the Elementary and Secondary Education Act of 1965 and creating unprecedented federal influence on SEAs and LEAs (Jeynes, 2007; McGuinn, 2006). NCLB sought to create systematic school improvement by allowing administrative discretion in the use of federal funds and experiments in administrative autonomy in exchange for state and district accountability for students' performance on standardized tests, with personnel and institutional consequences for failing to meet improvement targets (McDermott, 2011; McGuinn, 2006). Scholars Peterson and West hailed these efforts as an opportunity to "redirect educational thinking along new channels" (*No child left behind?*, 2003, p. 1), while

¹ See Appendices B and C for a flowchart of RF grant and sub-grant processes.

² See Appendix A for a brief overview of NCLB provisions.

then-secretary of education Paige (2003) and scholars such as Hewitt (2011) held that NCLB provided an opportunity to close racial achievement gaps.

The logic of NCLB's reforms was that by setting clear standards for schools to meet and imposing consequences for failure to make progress towards those standards, policymakers' "coercive accountability" would compel schools' improvement (Elmore, 2004; McDermott, 2011; McGuinn, 2006; Hess in *No child left behind?*, 2003, p. 58). In order to accomplish this goal, NCLB required states to create content and performance standards, assess students' performance towards those standards with standardized tests, and apply consequences to schools for failure to meet performance targets, or "adequate yearly progress" (AYP) ("NCLB," 2002, 20 USC 6311 Sec. 1111 (b) (2) (B)). With AYP pressures applied to motivate school improvement, RF served as a model and means of promoting the "scientifically based" instruction that was hoped would ameliorate racial achievement gaps.

This Study of RF. This study synthesizes research to analyze the mandates, implementation, and outcomes of the RF program. Information on RF mandates was collected through the USDOE website and reviews of NCLB legislation. Information on RF implementation was obtained by collection of the guidance documents published by USDOE, scholarly literature, and journalism on RF. Information on the outcomes of RF programs were collected through a review of government studies, independent assessments, and the scholarly literature. Relevant documents were obtained by searching the ERIC and Education Source databases. The term "*reading first*" was entered in conjunction with *mandate*, *guidance*, *implementation*, *impact*, and *outcome*. While the emphasis of this study was on the original research investigating RF, other pieces of scholarly work such as expert commentaries, literature reviews, and research syntheses were employed as anchors for this necessarily inter-disciplinary

analysis. For example, Edmonds and colleagues' (2009) synthesis of research anchored understanding of instruction's effect on secondary students' reading comprehension scores, while Corey, Phelps, Ball, Demonte, & Harrison's (2012) study provided background on the significance of instructional time.

Description of RF research topics. Fifty-three pieces of original research are included in this review. Twenty-four investigated features of both the implementation and outcomes of RF, the majority (46 or 88%) investigated features of RF implementation, and fewer (32 or 62%) investigated the impact of RF on measures of student achievement. Nineteen studies evaluated curricula or tests involved in RF, nine explored the importance of fidelity to RF's reading programs, and two investigated RF's effects on instructional time.

Methods used in RF research. Seventeen of the studies reviewed employed both quantitative and qualitative methods, with seventeen employing only quantitative and fifteen employing only qualitative methods. Of the 46 that studied features of RF implementation, 38 employed qualitative methods and 32 employed quantitative methods. Fifteen of the 32 quantitative studies employed surveys. Only one study (Gersten, Dimino, Jayanthi, Kim, & Santoro, 2010) employed fully experimental design to assess the outcomes of RF. Few comparison schools were available because RF had a national effect and targeted low-income schools with students who struggled to read and worked to disseminate program features, preventing much matched comparison analysis to schools that did not use RF practices (e.g. B. Gamse, Boulay, Rulf-Fountain, Unlu, & Society for Research on Educational Effectiveness (SREE), 2011). As a result, studies were susceptible to sampling and cohort biases. Five studies reviewed (Alvermann et al., 2007; Callow-Heusser & Chapman, 2007; B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst,

Boulay, Unlu, et al., 2008; Miller et al., 2007) attempted to overcome this difficulty by matching schools by background characteristics and employing regression discontinuity to discriminate program effects, and one employed difference-in-difference models (Moss et al., 2006). When assessing students' scores on measures of reading comprehension as a result of RF, 27 studies employed state tests or DIBELS. Of these studies, eleven triangulated these results with those of other tests, three employing the SAT-10, two the ITBS, two the TerraNova, and one using the PPVT-III. Ten studies assessed the outcomes of RF with measures other than state tests or DIBELS.

Mandates of Reading First

As established under the NCLB Act, RF was designed “to provide assistance to State educational agencies and local educational agencies in:” “establishing reading programs for students in kindergarten through grade 3 . . . based on scientifically based reading research;” “preparing teachers . . . through professional development and other support;” “selecting or administering . . . reading assessments;” “selecting or developing effective instructional materials . . . programs, learning systems, and strategies;” and “strengthen[ing] coordination among schools, early literacy programs, and family literacy programs” in order to “improve reading achievement for all children” (USDOE, 2002, sec. 1201). The program aspired to “implement methods that have been proven to prevent or remediate reading failure within a State,” “ensur[ing] that every student can read at grade level or above not later than the end of grade 3” (USDOE, 2002, sec. 1201).

The design employed to create this large-scale change in reading instruction was formula grants³. State educational agencies seeking a grant would submit an application to the U.S>

³ See Appendices B and C for a flowchart of RF grant and sub-grant processes.

Secretary of Education for grants that would last up to six years. States received RF monies based on the number of children ages 5-17 who live in the state as a part of a family with income below the poverty line (Garcia & Wiese in Morrow et al., 2009). States could reserve 20% of funds for state-level RF activities, including professional development, technical assistance to LEAs, and administration of sub-grants. The remaining 80% would be allocated to local educational agencies through competitive sub-grants that targeted low-literacy and high-poverty areas and provided reading programs, assessments, and professional development. Funds appropriated for RF by Congress hovered around \$1 billion per year, starting at \$900 million in 2002, peaking at \$1,041 million in 2005 (USDOE, 2009). Appropriations were slashed 61% in 2008 to \$393 million (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009; Glenn & Brainard, 2007) amid revelations of USDOE misconduct (Ashby & General Accounting Office, 2007; USDOE, Office of Inspector General, 2006, 2007a, 2007b) and mixed reviews in evaluation studies (e.g. B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2008). The program was closed the following year.

Federal Requirements. Approval of applications was to be conducted by the U.S. Secretary of Education, with the recommendation of a peer review panel (USDOE, Office of Inspector General, 2006). This panel was required to include “three individuals selected by the [USDOE] Secretary,” “three individuals selected by the National Institute for Literacy,” three individuals selected by the National Research Council of the National Academy of Sciences,” and “three individuals selected by the National Institute of Child Health and Human Development” [USDOE, 2002, sec. 1203 (c) (2) (A)]. Members of the review panel were legally required to include individuals with the expertise necessary to evaluate applications, individuals

who provide professional development for instruction based on “scientifically based reading research” (SBRR), and individuals who have expertise in reading assessments [USDOE, 2002, sec. 1203 (c) (2) (B)]. States whose applications were accepted were provided grant monies by a formula based on the number of children ages 5-17 living in households below the poverty line in each state (Morrow et al., 2009).

Targeted Assistance Grants. Additional targeted assistance grants were required to provide LEAs in need additional resources in states that demonstrated an increase in third graders in demographic subgroups “reaching the proficient level in reading” and that schools were improving reading skills of student in grades one, two, and three across two consecutive years (USDOE, 2002, sec. 1204). Targeted assistance grant monies were disbursed according to a formula based on the number of school-age children living in households below the poverty line in a given state [USDOE, 2002, sec. 1204 (c) (1)]. Targeted assistance applications were to be assessed by USDOE, and continuing grants made on the basis of whether states were able to demonstrate growth over the preceding two years (USDOE, 2002).

Federal Program Support. The U.S. Secretary of Education was allowed to use RF funds to provide technical assistance to SEAs, LEAs, and schools that requested it (USDOE, 2002, sec. 1206). The National Institute for Literacy was required to disseminate information about: “scientifically based reading research;” “schools, local educational agencies, and State educational agencies that have effectively developed and implemented classroom reading programs . . . that have been identified as effective through evaluation and peer review;” and reading programs that contain the essential components of reading instruction as supported by scientifically based reading research” to “recipients of Federal financial assistance under [title I], title III, the Head Start Act, The Individuals with Disabilities Education Act, . . . the Adult

Education and Family Literacy Act;” and schools funded by the Bureau of Indian Affairs (USDOE, 2002, sec. 1207).

Required Evaluations. The U.S. Secretary of Education was required to evaluate the impact of RF on the referral to, eligibility for, and special education services provided under the Individuals with Disabilities Education Act (USDOE, 2002, sec. 1206). USDOE was required to contract for an external evaluation of RF by an independent organization spanning five years and employing a “rigorous, scientifically valid, quantitative” method [USDOE, 2002 sec. 1205 (a)]. This evaluation was required to include analyses or measurements of: “the relationship between each of the essential components of reading instruction and overall reading proficiency”; “whether tools used by State educational agencies and local educational agencies measure essential components of reading”; “how State reading standards correlate with the essential components of reading instruction”; whether the receipt of targeted assistance grants . . . results in an increase in the number of children who read proficiently”; “the extent to which specific instructional materials improve reading proficiency”; “the extent to which specific . . . reading assessments assist teachers in identifying specific reading deficiencies”; “the extent to which professional development programs implemented by State educational agencies [under RF] improve reading instruction”; “how well students entering the teaching profession are prepared to teach the essential components of reading instruction”; changes in students’ interest in reading and time spent reading outside school”; and “any other analysis or measurement . . . determined to be appropriate by the Secretary” [USDOE, 2002, sec. 1205 (c)].

State Requirements. State grants were awarded based on their applications. Applications included provisions for: program review by USDOE; description of state plans to assist LEAs in choosing appropriate reading assessments and programs; description of state plans for effective

coordination of federal and state funds for professional development to improve instructional practices based on SBRR; how state sub-granting processes meet legal requirements; how states will assess RF compliance; and how states will build on and promote coordination with other literacy initiatives [USDOE, 2002, sec. 1203 (b)]. Grants were required to be made “in relation” to the number of economically disadvantaged students and students with reading difficulties in a district [USDOE, 2002, sec. 1204 (c) (1)]. In addition, states were to describe the criteria they would employ in distributing sub-grants to LEAs [USDOE, 2002, sec. 1204 (c) (3) (C)] and provide notice to LEAs of the availability of competitive sub-grants and the means for applying for them [USDOE, 2002, sec. 1204 (d) (2)].

In order to obtain targeted assistance grants released in fiscal year 2004, states were required to submit an application that detailed requested funding and the criteria to be used in distributing sub-grants [USDOE, 2002, sec. 1204 (c) (3) (C)]. States also had to demonstrate that an increasing percentage of third graders in demographic subgroups in schools served by grants were meeting proficiency levels and that schools served by grants were increasing the reading skill of students in grades 1-3 based on reading assessments over each two-year period [USDOE, 2002, sec. 1204 (a)]. States that received targeted assistance grants were required to notify all LEAs of the opportunity to apply for sub-grants, and to evaluate applications through a competitive process based on LEAs’ relative need [USDOE, 2002, sec. 1204 (d)].

Local Requirements. In order to obtain sub-grants, states would require local educational agencies to apply and make awards: in relation “to the number or percentage of students in kindergarten through grade 3 who are reading below grade level”; to schools “identified for school improvement measures” under NCLB adequate yearly progress provisions; and that “have the highest percentages or numbers of children” targeted by Title I based on a

formula using measures of poverty [USDOE, 2002, sec. 1202 (c) (6) (A-B)]. LEAs must demonstrate progress in any two consecutive years in improving the percentage of third graders from economically disadvantaged, racial and ethnic minority, and disability subgroups rated proficient in reading in RF schools, as well as reading skill increases for students in grades 1-3 in RF schools [USDOE, 2002, sec. 1204 (d) (4) (B)].

Local sub-grant funds would be used to select, purchase, and carry out reading assessments, instructional programs, instructional materials, professional development to teachers of students in kindergarten through grade 3, collection and summarization of data for accountability and instructional improvement, and promotion of reading programs that provide students access to reading material [USDOE, 2002, sec. 1202 (c) (7) (A)]. The reading programs employed by local educational agencies were required to employ the “essential components of reading instruction” [USDOE, 2002, sec. 1202 (c) (7) (A)], defined by the National Reading Panel as phonemic awareness, phonics, vocabulary, fluency, and comprehension (National Institute of Child Health and Human Development, 2000; Scott & Fagan, 2005) and to be “based on scientifically based reading research,” including screening to “identify students who may be at risk for reading failure or who are having difficulty reading” [USDOE, 2002, sec. 1202 (c) (7) (A) (ii & iv)].

Implementation of Reading First

In the words of Moss et al. (2006), RF seemed to have been “implemented in schools and classrooms as intended by the legislation” (p. 14). A majority of educators surveyed expressed satisfaction with the program, and with USDOE’s assistance (Ashby & General Accounting Office, 2007; Scott, 2007b). Mixed and difficult-to-interpret results in Gamse, Jacob, Horst, Boulay, and Unlu's (2008) impact study and revelations of USDOE misconduct (Ashby &

General Accounting Office, 2007; Manzo, 2002, 2005, 2007a; USDOE, Office of Inspector General, 2006, 2007a, 2007b), however, preceded a 61% cut in funding, and eventually the program's demise (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009).

USDOE Implementation of RF. RF programs were established in all 50 states and the District of Columbia, in 2003 reaching 1,415 districts and 4,774 schools (Scott & Fagan, 2005) and growing to 1,809 districts and 5,880 schools in April 2007 (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b). Appropriations for RF totaled over \$6 billion from the program's rollout in 2002 through its attenuation in 2008 and termination thereafter (USDOE, Office of Inspector General, 2007a).

USDOE officials and contractors assessed state applications and provided technical assistance on the implementation of RF programs to states through regional centers, offering guidance documents, workshops, and feedback during applications and monitoring visits (Ashby & General Accounting Office, 2007). Reporting in October 2002 indicated that 29 of 40 states applying for RF funds were asked to adjust their application and re-apply (Manzo, 2002), with 27 states reporting they were required to adjust their application three or more times (Ashby & General Accounting Office, 2007).

USDOE Guidance. USDOE helped to disseminate information about RF by employing guidance documents, presenting workshops, and providing feedback to SEAs and LEAs during application and monitoring visits (Ashby & General Accounting Office, 2007). Technical assistance centers were established to help states and districts align their programs with NCLB SBRR requirements (McCallion, 2008). States largely indicated USDOE guidance was useful, with over 40 states agreeing it was adequate to address key application features, and with disagreement from eight states (Ashby & General Accounting Office, 2007). USDOE's guidance

practices enforced their view of SBRR-aligned programs (Ashby & General Accounting Office, 2007; Manzo, 2005, 2007a), which drew on the work of the NRP (National Institute of Child Health and Human Development, 2000) used to construct the law (Cunningham, 2001).

USDOE Representation of RF. An April 2002 guidance document (USDOE, Office of Elementary and Secondary Education, 2002) explained that the purpose of RF was to “help States and districts apply” the available “scientifically based research,” promising to support “proven methods of early reading instruction” (p. 1). This document argued that “scientifically based reading research has identified five essential components of reading instruction” and that “programs funded under reading first will have to demonstrate their ability to address these components in a comprehensive and effective manner” (USDOE, Office of Elementary and Secondary Education, 2002, p. 2). The definition of scientifically based reading research articulated in the document emphasizes that research should be based on “empirical methods” of “observation or experiment” that allow “convincing documentation that the observed results were the product of intervention” (USDOE, Office of Elementary and Secondary Education, 2002, p. 4).

USDOE Technical Assistance. USDOE workshops emphasized the importance of teachers’ fidelity to the reading programs adopted under grants as key to RF’s success (Brenner & Hiebert, 2010). Workshops alluded to specific reading programs, which states interpreted as the promotion of specific reading curricula, expressly prohibited under NCLB (USDOE, Office of Inspector General, 2007a). Subsequent reports found that USDOE failed to establish adequate safeguards for its officials or representatives to forestall promotion of particular curricular materials (USDOE, Office of Inspector General, 2007a). USDOE also failed to address issues of bias in approving individuals to provide RF technical assistance, and in employing several

individuals with financial ties to commercial reading programs (USDOE, Office of Inspector General, 2007a). Additional conflicts of interest were present among USDOE technical assistance contractors, who provided incorrect information to LEAs (USDOE, Office of Inspector General, 2007b). Referrals of LEAs to the Oregon RF Center and the Florida Center for Reading Research reviews by RMC Research Corporation, a federal contractor for technical assistance, may have contributed to some SEA perceptions that there was an approved list of reading programs (USDOE, Office of Inspector General, 2007b).

USDOE Requirements for RF Programs. USDOE established a variety of requirements for RF programs. RF programs across the country were required to employ a comprehensive reading program, and use grouping strategies including small-group instruction based on ongoing assessment measures (Morrow et al., 2009; USDOE, Office of Elementary and Secondary Education, 2002).

Programs were required to integrate the five essential components of reading instruction—phonemic awareness, phonics, vocabulary, reading fluency, and reading comprehension—into explicit, coherent instruction with adequate opportunities for practice during a 90-minute uninterrupted daily block (Healy, 2007; Morrow et al., 2009; Stewart, 2004; USDOE, Office of Elementary and Secondary Education, 2002). Reading programs were required to employ small groups organized by students’ assessment data, with instruction appropriate to groups’ needs (USDOE, Office of Elementary and Secondary Education, 2002; USDOE, Office of Inspector General, 2006). Programs were required to assess students’ progress in the five essential components and gauge student needs, using relevant assessments three times a year and disaggregating by income, race, English proficiency, and disability status at school, district, and state levels (Morrow et al., 2009; USDOE, 2002).

In order to support RF programs' success, districts were required to use RF funds to hire a reading coach to support instructional improvement, help teachers to collaborate by creating grade-level meetings, use data from regular student assessments to make instructional decisions, and employ the RTI intervention model for struggling readers (Chapman, 2010; Nelsestuen et al., 2009). Districts were required to report students' performance to the state, which enforced RF program adherence through site visits and monitoring, and states were required to report to USDOE, who enforced state-level RF program requirements (Healy, 2007). USDOE required that RF programs use funds for reading programs, materials, remedial programs, ongoing monitoring of student progress, and educator professional development necessary to support effective program implementation in order to receive grant monies (McCallion, 2008; USDOE, Office of Elementary and Secondary Education, 2002).

USDOE Treatment of State Programs. USDOE required that state applications include the procedures used to ensure that professional development activities and LEAs' instructional programs met requirements for basis in SBRR (USDOE, Office of Elementary and Secondary Education, 2002). USDOE provided guidance to states through formal guidance and technical assistance documents, as well as through informal channels of consultants, asking 33 states to address issues in states' use of instructional assessments and 25 states to address instructional strategies and programs. Forty-eight states reported needing to modify their application once, while 27 reported modifying their application three or more times (Ashby & General Accounting Office, 2007). Some of this confusion was a result of the failure of USDOE to follow its own rules: rather than provide the response of the expert review panel to state applications, the U.S. Secretary of Education provided an edited version of their comments as a report, which including altering meaning (Ashby & General Accounting Office, 2007). USDOE officials and contractors

promoted their conception of SBRR programs under RF, stacking expert review panels with those who favored their philosophical approach, and in some cases suggesting explicit programs for LEAs to adopt in violation of NCLB provisions (Ashby & General Accounting Office, 2007; Manzo, 2007a).

The expert review panel USDOE assembled to make recommendations for state RF programs did not meet legally established guidelines, which would have called for selection of experts by outside organizations and an oversight panel for appeals (USDOE, Office of Inspector General, 2006). The Office of Inspector General's report on RF mismanagement found that expert review panels were illegally selected predominantly by USDOE rather than the organizations designated in the act, and no oversight panel for RF applications was ever created (McCallion, 2008; USDOE, Office of Inspector General, 2006). The experts selected were screened for conflicts of interest, but USDOE inspector general's reports found that this process was inadequate, and panels featured experts with strong financial ties to textbooks and programs subsequently approved for use in RF curricula (Ashby & General Accounting Office, 2007; Healy, 2007; Manzo, 2007a; USDOE, Office of Inspector General, 2006, 2006). In one example of failure to address conflicts of interest, Edward Kame'enui and Deborah Simmons, authors of *A Consumer's Guide*, a checklist used in many states to evaluate core reading programs, were employed as contractors to provide technical assistance for RF to many states (Grunwald, 2006; Manzo, 2004, 2005).

USDOE mismanagement contributed to states' difficulties in the application process, in some cases resulting in violation of NCLB requirements (Ashby & General Accounting Office, 2007). The review panel's comments on state applications were not provided to states as suggested in USDOE guidance (USDOE, Office of Elementary and Secondary Education, 2002).

Instead, the U.S. Secretary of Education summarized the expert review panel's comments in a report, with significant editing that altered the meaning of expert comments (USDOE, Office of Inspector General, 2006). Several state officials reported that technical assistance from USDOE was vague and directed them to a University of Oregon review of assessments (where contractors Edward Kame'enui and Deborah Simmons were affiliated and their work *Consumer's Guide* posted), while more had re-submissions that clearly did not incorporate panelists' comments because RF director Chris Doherty's reports failed to adequately summarize them. (USDOE, Office of Inspector General, 2006).

A 2007 report by the General Accounting Office (2007) found that USDOE failed to establish written procedures that would guard against federal officials or contractors prompting districts to adopt or discontinue particular reading products. USDOE officials or contractor representatives were reported by four states to suggest adoption of specific programs or assessments, and ten states reported that USDOE representatives suggested they eliminate specific programs and assessments (Ashby & General Accounting Office, 2007). Doherty engaged in conference calls with states regarding their applications rather than the review panel, and internal emails sent by the director suggest that "remarks to groups . . . or face-to-face meetings" provide "opportunities for . . . extralegal requirements" (USDOE, Office of Inspector General, 2006). Officials worked to prevent states from using the Reading Recovery program in their RF grant plans (Manzo, 2007a, 2007c). Wisconsin educators reported pressure to abandon their existing literacy efforts in favor of commercial products (Manzo, 2006a). In some cases, this pressure continued after USDOE had approved state RF plans (USDOE, 2006). Some RF applicants were funded without documentation that they met expert review panel requirements (USDOE, Office of Inspector General, 2006).

USDOE failed to employ written, transparent procedures in annual state implementation monitoring, causing confusion in monitoring procedures and timelines among state officials seeking to meet them (Ashby & General Accounting Office, 2007). In monitoring reports, state officials felt they were assessed on whether programs implemented by districts met USDOE's vision of SBRR rather than federally-approved state standards for RF programs (Ashby & General Accounting Office, 2007). In four different incidents, states' reported that USDOE officials or regional technical assistance center staff contacted state officials about district RF applications (Ashby & General Accounting Office, 2007).

Curricula Approved. The curriculum plans approved by USDOE under RF were nominally based on the NRP analysis of what constitutes SBRR. Bell's (2003) review of RF grant recipients; however, found discrepancies between the NRP's conclusions and the programs approved by USDOE under RF. *Consumer's Guide* was approved for use at state and/or local levels in 42 states (Scott & Fagan, 2005), suggesting that USDOE agreed with its outright endorsement of decodable texts, a topic on which the NRP report remained ambivalent (Bell, 2003; Cunningham, 2001; National Institute of Child Health and Human Development, 2000).

Most of our knowledge of curricula approved under RF comes from government evaluations (Ashby & General Accounting Office, 2007; Bell, 2003; B. Gamse et al., 2011; B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006, 2008; USDOE, Office of Inspector General, 2006). The Center on Education Policy's exhaustive surveys of state RF officials and nationally representative samples of teachers about NCLB provide corroborating evidence (Scott & Fagan, 2005; Scott, 2006, 2007a, 2007b). Bell's (2003) review of thirteen states' RF grant applications provides some accountability to the USDOE approval process. A variety of smaller

studies provide powerful evidence of the curricula adopted on the district level (Bonds, 2010; Brown-Wright et al., 2013; Elias, 2009; Gerstl-Pepin & Woodside-Jiron, 2005). Appendix D provides a chart from Ashby & General Accounting Office' (2007) report summarizing some of the most widely-adopted curriculum products under RF.

USDOE's decisions in approving curricula affected what programs states employed, sometimes making controversial endorsements. Bell's (2003) review of thirteen states' RF grant applications found no federally-mandated list of approved reading programs common across states. RF did, however, set horizons for programming and appropriate materials, with state applications demonstrating commonalities (Manzo, 2004). Twenty-nine states allowed districts to determine which reading programs met SBRR requirements with state oversight, while 22 created a state-approved list of reading programs (Ashby & General Accounting Office, 2007). Among these 22, fifteen reported using *Consumer's Guide* to select reading programs, and Houghton Mifflin, McGraw-Hill, and Harcourt were most frequently represented on these lists, with GAO estimates that suggested 11-23% of participating schools employed these programs (Ashby & General Accounting Office, 2007). Four of the five most cited core reading programs in RF schools were also cited among the five most popular in Title I schools (Harcourt Trophies, McGraw-Hill Open Court, Scott Foresman Reading, and Houghton Mifflin Reading) (Moss et al., 2008). Later What Works Clearinghouse assessments found insufficient evidence to include McGraw-Hill's Open Court, Scott Foresman Reading, and Houghton Mifflin Reading, and other approved programs in their listing of programs with sufficiently rigorous review (Institute of Education Sciences (ED), 2007). Certain programs, including *Success for All*, *Reading Recovery*, *Kaplan SpellRead*, and *Start Making a Reader Today* were barred from or employed much less among RF programs

than among schools generally, despite being rated by independent organizations as research-based (Healy, 2007; Institute of Education Sciences (ED), 2007; Manzo, 2007b, 2007c).

RF grant applications widely employed the DIBELS standardized test to evaluate students' performance (Li & Zhang, 2008), and USDOE officials were reported to have steered states toward it (Manzo, 2005; *The Routledge international handbook of English, language and literacy teaching*, 2010). In 2005, Center for Education Policy reports found that 37 of the 50 states participating in RF required the use of DIBELS in LEAs' assessments, and five additional states included it among a list of options (Scott & Fagan, 2005), while government reports suggest that 45 states employed DIBELS in their RF programs (Ashby & General Accounting Office, 2007). The test has garnered controversy, drawing debate over its efficacy at predicting reading comprehension and ability to provide diagnostic information to inform instruction (e.g. Li & Zhang, 2008; Morrow et al., 2009; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Shelton, Altwerger, & Jordan, 2009).

State Implementation of RF

State implementation of RF included different models under which LEAs could adopt programs. In 29 states, LEAs were responsible for determining which reading programs met SBRR requirements with SEA oversight, while 22 states created lists of reading programs or materials approved by the state as meeting SBRR requirements (Ashby & General Accounting Office, 2007). Seventy-six percent of states with RF grants reported coordination with Title I, which may have been a pathway for dissemination of RF features to Title I schools in the absence of formal grants (Scott, 2007b). Prominent in state RF implementation were offers of professional development, which 48 states reported employing to assist districts in their efforts (Scott, 2007a).

In some cases, states reported difficulty with their role in RF. In Ashby & General Accounting Office's (2007) survey of state RF directors, one state official reported feeling ill-equipped to apply SBRR principles in evaluating reading programs, and relied on other states' reviews and lists of reading programs to determine which to allow. *Education Week* reporting suggests that states whose applications had been denied reported adopting DIBELS as an assessment and *Consumer's Guide* as a checklist for evaluating core reading programs. This may be the result of state confusion during the application process, or it could signal the difficulty for states and districts in sifting accurate SBRR advice from the few empirical studies available and the many resources purporting to offer SBRR advice (Foorman & Nixon, 2006; McCallion, 2008). Twenty-two states reported that it was difficult for them to assist RF schools that had not demonstrated the progress on reading scores required under RF (Ashby & General Accounting Office, 2007). After alleged attempts by USDOE officials and consultants to pressure Wisconsin educators to abandon their existing literacy efforts in favor of commercial products favored under RF, Madison, Wisconsin superintendent Art Rainwater returned RF grant monies totaling some \$2 million (Manzo, 2006a).

When Gamse and colleagues (2011) surveyed states as to what features of RF they would sustain after the drastic attenuation in funds in 2008, states reported they would be able to sustain some services only if RF program elements were included in new state standards or literacy initiatives. Sixty-seven percent of states anticipated they would require additional funding to support RF-related activities after grant funding fell: 39% of RF directors reported it was unrealistic to continue RF programs using other state funds due to widespread financial shortfalls (B. Gamse et al., 2011).

Local Implementation of RF

Methodological Concerns. Knowledge of local implementation of RF comes primarily from government surveys of teachers, principals and reading coaches that compare RF and Title I schools (Moss et al., 2006, 2008). Additional information was provided by the Center for Education Policy's surveys of state RF directors and nationally representative samples of teachers that sought to examine changes resulting from NCLB (Scott & Fagan, 2005; Scott, 2006, 2007a, 2007b). The surveys that serve as the primary source of information on implementation are susceptible to social desirability bias, where respondents are more likely to report favorable outcomes. If respondents knew surveyors were examining for features of RF, they may have overrepresented scores for questions relating to features of RF, making implementation appear more successful than it was. Several qualitative studies provide school-level perspective on implementation (Elish-Piper & L'Allier, 2011; Gerstl-Pepin & Woodside-Jiron, 2005; Kersten & Pardo, 2007).

RF's Influence. As of April 2007, RF provided grant funding directly to 1,809 districts (7%-13% of the nation), who distributed it to 5,880 schools (about 6% of the nation; Ashby & General Accounting Office, 2007; B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; Scott, 2007b). In addition to this wide influence, RF also required change among 67% of a nationally-representative sample of districts (Scott, 2007b). There are several ways in which RF influence is likely to be greater than these figures. Eighty percent of schools and 76% of districts in Scott' (2006) survey reported coordinating Title I and RF funding, and many described changing Title I instruction to match RF and expanding RF programs to non-RF schools. Four of the five most cited core reading programs in RF schools were also among the five most popular in Title I schools (i.e. Harcourt Trophies, McGraw-Hill Open Court, Scott Foresman Reading, and Houghton Mifflin Reading; Moss et al., 2008). Half of participating

districts in (Scott, 2007b) survey reported implementing features of RF in higher grades, as did some Bureau of Indian Affairs schools (Callow-Heusser & Chapman, 2007). With 42 states responding, estimates suggest 3,274 non-RF districts participated in RF professional development (Scott, 2007b). Some states may have adopted RF program features as “best practices.” In Missouri, only 10% of principals in (Powell, Higgins, Aram, & Freed's (2009) survey reported receiving a RF grant, but 36% reported having a literacy or reading coach in spring of 2006.

Character of Changes. Comparison of RF schools to matched Title I schools using surveys and difference-in-difference models found that RF schools devoted more time to reading instruction, and reported greater teacher knowledge of SBRR-aligned materials and practices (Moss et al., 2008). Additionally, RF schools were more likely to: have reading coaches who assisted in reading program implementation, use SBRR-aligned reading materials, use assessments to guide instruction, provide struggling readers intervention services, and provide staff professional development in the NRP's essential components of reading instruction (Moss et al., 2008). In Scott's (2006) survey, state RF directors reported “adding or modifying reading assessments, putting more emphasis on the five components of reading advocated by Reading First, and providing teachers with professional development in reading” (p. 4).

Instructional Materials. When asked to report changes in response to RF, schools most often reported purchasing new textbooks or reading materials that fit SBRR requirements, and did so at more than twice the rate of Title I schools (39% of RF schools v. 16% of Title I; Moss et al., 2006; Scott, 2006). In the years following implementation, RF schools were more likely than Title I schools to add new supplementary materials and adopt new materials for English Language Learners (Moss et al., 2006, 2008). In comparisons with Title I schools, however, RF

teachers reported lower levels of experience with supplemental reading materials (Moss et al., 2006), and no difference in the availability of differentiated instructional materials for struggling readers was found in subsequent survey (Moss et al., 2008).

90 min. Instructional Block. RF schools often reported changing the time allotted for reading instruction in response to RF mandates (Scott, 2006). RF schools were more likely than Title I schools to report having a 90 minute uninterrupted block for reading instruction, reporting 103 minutes spent on reading daily compared to 81 min. in Title I schools (Moss et al., 2006, 2008)

Use of Assessments. No difference was reported between RF and Title I schools in the rate at which assessments were employed (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006). RF teachers were, however, more likely than Title I teachers to report regularly scheduled, formal time set aside to use assessment data to plan instruction, and were more likely to employ assessments to identify struggling readers (Moss et al., 2006, 2008). Teachers in RF schools reported using assessments for grouping students and identifying students who needed additional interventions (Scott, 2007b), and were more likely to do so than their Title I counterparts (Moss et al., 2008).

Intervention Programs. Initial reports from RF schools indicated they were more likely to add intervention programs for struggling readers than their Title I counterparts (74% v. 43%; Moss et al., 2006), but in the 2006-07 school year 40% of both RF and Title I schools reported adding intervention programs for struggling readers (Moss et al., 2008). RF teachers were more likely than their Title I counterparts to put struggling readers into intervention programs (80% v. 63%), and reported doing so at an average of one week after identification (Moss et al., 2006, 2008).

Professional Development. RF teachers were more likely to report SBRR-aligned professional development, reporting an average of 12.1 hours more annually and covering more of the NRP's five essential components (especially fluency; 91% v. 74%) than Title I teachers (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2008; Powell et al., 2009).

Administrative Adjustments. Schools made a variety of administrative adjustments to adapt to RF. Teachers in RF and Title I schools were equally likely to report a school-wide focus on reading instruction and alignment of that instruction with state standards (Moss et al., 2006). In both RF and Title I schools, districts were generally responsible for determining which assessments to use, while principals and reading coaches were responsible for interpreting assessment results (Moss et al., 2006). RF teachers also reported more time to use assessment data to plan instruction, to observe reading instruction in other classrooms, to collaborate on lesson planning and instruction, and to receive instruction from a reading coach (Moss et al., 2006). Title I schools were more likely to report having certified special education teachers provide recommendations to plan instruction for struggling readers, and RF teachers in K-2 were less likely to report time being set aside to coordinate reading instruction for special education students (Moss et al., 2006).

Reading Coaches. Many schools reported that they hired a reading coach in response to RF mandates (Scott, 2006). RF schools were more likely to report have a reading coach than Title I schools and reading coaches reported spending more time spent training teachers (Moss et al., 2006, 2008), while more teachers reported receiving coaching (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). Rather than imagine reading coaches served simply to transmit RF visions of best practice, qualitative studies suggest reading coaches have helped to adapt the prescribed curriculum to local visions of best practice and students' developmental needs (Elish-

Piper & L'Allier, 2011; Kersten & Pardo, 2007; Lindemer, 2005; Pease-Alvarez & Samway, 2008).

Teachers' Reactions. Qualitative study of RF implementation has found that teachers have worked to adapt the scripted reading curriculum provided to the needs of their students, in some cases working against the will of district supervisors (Pease-Alvarez & Samway, 2008). Some adaptation was supported by reading program publishers, and principals made exceptions for teachers they thought were strong enough to leave the script (Pease-Alvarez & Samway, 2008).

Outcomes of the RF Program

Methodological Concerns

Studies of RF Outcomes. Twenty-nine of the studies included in this review investigated the outcomes of RF programs on measures of student achievement, with 21 of those studies including some investigation of implementation. Federal government-contracted studies provided the greatest insight (B. Gamse et al., 2011; B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006), and are triangulated by Center on Education Policy studies (McMurrer, 2008; Scott & Fagan, 2005; Scott, 2006, 2007a, 2007a) and one independent study (Wong-Ratcliff, Powell, & Holland, 2010). A variety of studies provide information on state-level RF outcomes (Alvermann et al., 2007; Baker et al., 2011; Bean et al., 2010; Boise State University College of Education Center for School Improvement and Policy Studies, 2007; Chapman, 2010; Day-Meeks, 2011; Dole, Hosp, Nelson, & Hosp, 2010; Foorman, Petscher, Lefsky, & Toste, 2010; Hayden, Trainin, Javorsky, Murphy-Yagil, & Cook, 2008; Miller et al., 2007; Murphy, Trainin, Yagil, Javorsky, & Hayden, 2007; Pfannenstiel, Seltzer, & Yarnell, 2008; Trainin,

Javorsky, Murphy, & Wilson, 2009; Trainin & Wilson, 2010; Westat & Learning Point Associates, 2008). Other studies compare core reading programs to “literature-based instruction” (Arya et al., 2005); analyze the relationship between program fidelity and outcome measures (Bowers, 2011; McGill-Franzen, Zmach, Solic, & Zeig, 2006) or reading coaching and outcome measures (Elish-Piper & L’Allier, 2011); assess the performance of a professional development program (Gersten et al., 2010); analyze the efficacy of DIBELS at predicting other measures of reading comprehension (Shelton et al., 2009); examine the use of federal funds under NCLB (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009); or use multiple measures to compare the effectiveness of two core reading programs (Skindrud & Gersten, 2006). Additional information on the outcomes of RF programs was obtained using qualitative studies (Elish-Piper & L’Allier, 2011; Gerstl-Pepin & Woodside-Jiron, 2005; Kersten & Pardo, 2007).

Methods Employed. Only one of the studies reviewed here employed a fully experimental design (i.e. Gersten et al., 2010). Many employed a quasi-experimental design due to the difficulty of finding a comparable control group for a program that had a nationwide effect and targeted the highest-poverty schools, which left many designs open to sampling and cohort biases. One study employed difference-in-difference models to negotiate this challenge (Moss et al., 2006). Three studies (Callow-Heusser & Chapman, 2007; B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008) employed a regression discontinuity analysis to address these issues, although criticism has been leveled that these studies are not valid comparisons because they do not use an uncontaminated control group (Baker et al., 2011; Herlihy, Kemple, Bloom, Zhu, & Berlin, 2009; Mitchell et al., 2008). Sixteen studies employed state tests or DIBELS to measure student outcomes. Such measures should be treated with caution. State tests are susceptible to score

inflation, whose influence is difficult to measure, and which may corrupt measures of students' skills (Koretz, 2008). DIBELS samples decoding and phonemic awareness skills heavily, and may reveal more about effective RF implementation than students' increasing skills in gaining meaning from what they read (Arya et al., 2005; Shelton et al., 2009). Such measures additionally made interstate comparison difficult, and caused delays and calls for technical support in a congressional commission asked to analyze states' results (Manzo, 2007d). Five studies triangulated state test or DIBELS results with other tests, while ten assessed the outcomes of RF with other measures only. Center on Education Policy studies, which predominantly relied on survey information, should be treated with caution since they may be susceptible to social desirability bias, where respondents over-report responses that they think are socially desired.

Use of RF Funding

Chambers and colleagues' (2009) analysis of the use of 2004-05 school year funds under NCLB provides significant information about the use of RF funds by SEAs and LEAs. Of the \$1,204 million allotted for RF, approximately \$815 million or 72.3% went to school districts, 10.7% was not reported, and 0.2% went to other agencies (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009). Some \$677 million, or 83% of the grant monies awarded to districts, reached individual schools, which represented 2% of all schools nationally (compared to Title I's 56%) and 4% of elementary schools (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009). RF funds served 27% of all urban districts and 6% each of suburban and rural districts (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009).

Sixty-five to sixty-seven percent of RF funding was used for instruction in the 2004-05 school year, less than the 72-75% of Title I and III funds, with \$188 million in RF funds going to

instructional materials and services (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, & Stullich, 2009). RF grant monies contributed \$103-106 million to professional development, amounting to 6% of all ESEA professional development funding. Forty-two percent of RF funds, or some \$342 million, were spent on instructional staff.

Effect on Curricula

Perceived Change. A variety of changes as a result of RF have been captured by surveys. Sixty-nine percent of states reported great or very great improvements in the teaching of reading since implementing RF (Ashby & General Accounting Office, 2007), while a majority of the states surveyed in Scott's (2006) report indicated that RF grants were an important or very important cause of student achievement increases. RF teachers were likely to report covering more NRP essential components in the last month than Title I teachers, though no difference was found in reported availability of differentiated instructional materials for struggling learners or in teachers' use of assessments to inform classroom practice (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008).

Quality of Curricula. State, and slightly less often, district officials were likely to report that RF's instructional programs and assessments were important or effective measures at improving student achievement (Scott, 2006, 2007a, 2007b; Zhang & Center on Education Policy, 2008). Districts identified aligning curriculum and instruction with standards and assessments as effective at improving student achievement in Title I middle and high schools (Scott, 2007a). Many districts identified local policies as important or very important causes of ELA achievement (Scott, 2007a; Zhang & Center on Education Policy, 2008)

RF teachers were more likely to report that the assessments coordinated with their core reading program were most useful, while Title I teachers were more likely to report their

informal assessments were (Moss et al., 2006). RF teachers were more likely to report their curricula and instructional practice were SBRR-aligned, that they had access to reading intervention programs, and that high-quality instructional materials were available than their Title I counterparts, including new materials for ELLs (Moss et al., 2006). Reading coaches and teachers in RF schools were more likely to report teachers as knowledgeable about SBRR practices (Moss et al., 2008)

Quality of Administrative Changes. Likelihood of RF and Title I staff to report time set aside to coordinate with ELL staff grew from the 04-05 to 06-07 school years (Moss et al., 2008). Two-thirds of districts responding to Scott's (2007a) survey reported that district policies unrelated to NCLB were important to school improvement. Districts were likely to report that: increasing the use of student achievement data to inform instructional decisions; better quality or additional teacher professional development; extending the school day or year and improvement of school planning were effective strategies for educational improvement (Scott, 2007a). States were less likely to endorse the use of consultants or hiring full-time staff to support teacher development (Scott, 2007a). A majority of states reported that RF professional development improved students' achievement or was an effective strategy for school improvement (Ashby & General Accounting Office, 2007; Scott, 2007a, 2007b). When surveyed in 2008 as RF programs were facing funding cuts, state RF directors most often reported the use of literacy coaches would be worth sustaining (56%), followed by RF materials and curricula (39%), data-driven instruction (35%), use of assessments (35%), and SBRR-aligned instruction (33%) (B. Gamse et al., 2011)

Reading Instruction Time. Several studies provide insight into changes in the use of time under RF programs (Brenner & Hiebert, 2010; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006, 2008).

90 Minute Reading Block. In comparison to Title I schools, RF schools were more likely to have a 90 minute block of uninterrupted reading instruction, and spent 18.5 more minutes on reading instruction than Title I schools each day, averaging 105.7 minutes a day (Moss et al., 2006, 2008).

NRP Essential Components. RF implementation and evaluation studies found that RF schools spent moderately more instructional time addressing the NRP's essential components of reading instruction, providing highly explicit instruction, and providing opportunities for practice in those components (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). These effects were less significant for early award sites (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b), and systematically declined over time in grade two (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008).

Relation to Score Improvements. RF programs demonstrated statistically significant increases in the time spent on reading instruction, but only modest increases in decoding scores and no statistically significant effects on reading comprehension (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b). The number of years schools employed RF practices did not affect its impact on reading instruction outcomes (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b). One minute increases in the average time spent on reading instruction daily was associated with a 0.07 point increase in

students' scaled test scores (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b)

Relation to Other Instructional Features. Generous estimates of how much time students spent reading in core reading programs ranged from 11.3 to 27.1 minutes, averaging 16.7 minutes, which authors argued was insufficient to help struggling readers reach proficiency (Brenner & Hiebert, 2010). RF teachers reported more time allocated for struggling readers to practice, and third-grade RF teachers were likely to provide students extra practice in decoding and fluency (Moss et al., 2006). In classroom observations, RF classrooms demonstrated a negative effect on measures of student engagement with print (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b)

Activity Quality. Central to RF's impact is the tone it set for classroom reading instruction practice, but it is ambiguous whether the programs most frequently adopted were more appropriate to students' needs than other possibilities. In an assessment of RF program impacts, only composite survey measures of teachers' use of activities for struggling readers was correlated to a school's likelihood of scoring in the top quartile (Moss et al., 2008).

Core Reading Programs. *Open Court*, a McGraw-Hill reading program, was approved in twice as many state applications as the *Success for All* program (Ashby & General Accounting Office, 2007). In Skindrud and Gersten's (2006) analyses, it provided a "more consistent review of phonemic awareness, decoding, and reading fluency in grades K through 6; more specific reading comprehension and writing programs, grades K through 6; and more extensive use of decodable and authentic readers" (p.402), lending students generally a moderately better score on SAT-9 reading comprehension measures. For students at the bottom quartile of test takers, however, performance was not statistically different, which teachers attributed to *Success for*

All's effectiveness at accommodating student differences and developing students' social skills (Skindrud & Gersten, 2006). Arya and colleagues' (2005) comparison of second grade classrooms using *Reading Mastery* and *Open Court* reading programs to those using "literature-based instruction" (where phonics was taught incidentally) found no significant differences on phonics measures in the Woodcock-Johnson Psycho-Education Battery. Additionally, students in "literature-based instruction" demonstrated higher scores and fewer unacceptable guesses about unfamiliar words in researcher-developed miscue analyses. RF's often-scripted programs were found to demonstrate no difference to programs in Title I schools on measures of student engagement with print (Dudley-Marling, 2005; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008).

DIBELS. The DIBELS assessment was widely adopted under RF programs, and there is evidence that USDOE pressured states to use the assessment (Li & Zhang, 2008; Manzo, 2005). The test consists of a series of one-minute subtests that allow teachers to quickly obtain a measures of students' phonological awareness and reading fluency, making it relatively easy for teachers to administer (Li & Zhang, 2008). Studies suggest that DIBELS' Oral Reading Fluency measure is predictive of ITBS (Schilling, Carlisle, Scott, & Ji Zeng, 2007) and SAT-10 reading comprehension measures (Roehrig et al., 2008). It has been criticized, however, for failing to take into account cultural factors in literacy development, over-emphasizing speed at the cost of comprehension, and lacking diagnostic information that would inform instruction, which critics argue is better obtained with the GRA+DE or OSELA assessments (Li & Zhang, 2008; Morrow et al., 2009). DIBELS was not predictive of Shelton and colleagues' (2009) measures of reading comprehension, which were designed to be culturally sensitive and emphasize meaning constructed from text rather than phonological skills or fluency.

Effect on Students

State Test Results. State test scores and DIBELS measures experienced growth under RF faster than non-RF schools among the studies included in this review, although this did not attenuate the achievement gaps between demographic subgroups (Alvermann et al., 2007; Bean et al., 2010; Boise State University College of Education Center for School Improvement and Policy Studies, 2007; Chapman, 2010; Dole et al., 2010; Foorman et al., 2010; Hayden et al., 2008; McGill-Franzen et al., 2006; Miller et al., 2007; Moss et al., 2008; Murphy et al., 2007; Pfannenstiel et al., 2008; Skindrud & Gersten, 2006; Torgesen, 2009; Trainin et al., 2009; Trainin & Wilson, 2010; Westat & Learning Point Associates, 2008; Wong-Ratcliff et al., 2010). DIBELS generally demonstrated greater score gains than state tests or other measures of student achievement. State tests generally demonstrated greater growth in RF schools than in non-RF schools

Other Standardized Measures. Tests independent of NCLB's accountability pressures (i.e. not employed as state tests for AYP determination) are less likely to experience score inflation and over-represent students' reading comprehension skills (Koretz, 2008). Federal- and state-level RF professional development efforts included large numbers of non-RF schools, and many schools adopted RF program features, so no impact studies were able to compare RF schools to schools that did not use RF practices (Baker et al., 2011; Herlihy et al., 2009). Results on comparisons of schools using low-stakes tests, however, are less sanguine than other measures. In a national comparison, RF programs demonstrated only modest increases in SAT-10 decoding scores for first grade students in one school year and no statistically significant effects on reading comprehension (B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). No

relationship was found between the number of years students were exposed to RF instruction and reading comprehension scores, and no statistically significant variability by site or grade was demonstrated (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). A small positive correlation was observed between instructional time spent on the NRP's essential components of reading instruction on SAT-10 scores, but it was affected by the model and sample used to estimate its effects (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008)

Few state evaluations compared RF schools to non-RF schools, more frequently comparing RF progress across years of implementation. Kansas, Oregon, and Bureau of Indian Affairs RF schools demonstrated higher SAT-10 scores among schools with more years of RF implementation, and yearly growth for the four years measured, though growth on SAT-10 was smaller than that on state tests and DIBELS (Baker et al., 2011; Chapman, 2010; Pfannenstiel et al., 2008). Kansas RF SAT scores grew to the national norm over the four years studied (Pfannenstiel et al., 2008) Georgia and Idaho students' performance on the ITBS generally ranged from stagnation moderate losses (Alvermann et al., 2007; Boise State University College of Education Center for School Improvement and Policy Studies, 2007). RF students in Idaho demonstrated 4-8% more students achieving at the 40th percentile or higher than the mean (Boise State University College of Education Center for School Improvement and Policy Studies, 2007) Comparison of Georgia students' normal curve equivalent scores, mean scores, and scores by cohort demonstrated consistent, statistically significant gains only in spelling and measures of phonemic awareness (Alvermann et al., 2007). Comparisons with non-RF schools that could be matched based on English proficiency, number of students receiving free and reduced price lunch, and demographic information showed no statistically significant differences, though observed means were higher among non-RF schools on all subtests (Alvermann et al., 2007).

Wisconsin RF schools demonstrated greater TerraNova scores and growth than control schools on measures of phonemic awareness, but the program demonstrated no impact on composite, reading comprehension, or vocabulary scores (Miller et al., 2007). Ohio RF schools demonstrated some growth in TerraNova scores across grades, but 8% or fewer students achieved at the 40th or 50th percentiles than the mean (Westat & Learning Point Associates, 2008).

Alternative Measures. Several alternative measures allow us to contextualize standardized tests of RF program effects. Arya and colleagues' (2005) analysis of the effects of “literature-based instruction” in comparison to *Reading Mastery* and *Open Court* reading programs found that students performed better on researcher-developed miscue analyses, making fewer unacceptable guesses about words they didn't recognize. This suggests that students receiving literature-based instruction were more likely to comprehend the meaning of the text they were reading, and was not predicted by their DIBELS oral reading fluency scores (Arya et al., 2005). A case study employing narrative policy analysis found that RF's scripted programs and decodable texts undermined student engagement and prevented teachers from socializing students into literate practices (Gerstl-Pepin & Woodside-Jiron, 2005). RF teachers' use of assessments likely contributed to a higher rate of referring struggling readers to intervention services than Title I counterparts, suggesting students were more likely to receive additional instruction (Moss et al., 2008). McGill-Franzen and colleagues' (2006) investigation of RF programs' relation to student drop-out suggested that the emphasis on grade-level curriculum and whole-class instruction in the core reading programs they compared may have prevented struggling readers from making the progress that they needed to meet grade-level expectations.

Effect on Teachers

Changes associated with RF programs have created administrative changes that have the potential to support teachers' improvement, but teachers reported they needed additional professional development to provide effective interventions for struggling readers (Moss et al., 2006). RF teachers reported receiving more hours of professional development and covering more of the NRP's essential components than Title I peers (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006, 2008) RF teachers reported having more time to use assessment data to plan instruction, observe reading instruction in other classrooms, collaborate on lesson planning and instruction and to receive instruction from a reading coach (Moss et al., 2006). RF teachers reported more support in interpretation of assessment data, specialist help diagnosing student needs, and intervention help for individual students (Moss et al., 2008). Teachers reported pressure to maintain fidelity to reading programs (Brenner & Hiebert, 2010; Scott, 2007b), which was likely difficult as one review of core reading programs noted little guidance was provided on implementing the 100 pages of scripted lessons provided each week (Brenner & Hiebert 2010). RF teachers were also less likely than Title I teachers to report experience with supplemental and reading intervention materials, probably due to their novelty (Moss et al., 2006). In the resulting confusion, teachers and reading coaches in qualitative studies have adapted RF programs to local visions of best practice and students' developmental needs (Elish-Piper & L'Allier, 2011; Kersten & Pardo, 2007; Lindemer, 2005; Pease-Alvarez & Samway, 2008).

Reading Coaches. Study of reading coaching suggests that teachers are more likely to adopt instructional changes they learned about from reading coaching than from other sources (Coburn & Woulfin, 2012). RF teachers were more likely to report support from reading coaches and reported more support than their Title I counterparts, while reading coaches in RF schools

reported spending more time in that role (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006). RF reading coaches were more likely than Title I peers to report providing training or professional development, coaching staff on a range of topics, organizing professional development, facilitating grade level meetings, compiling reading assessment data, and coordinating reading assessments (Moss et al., 2006)

Chapter 3: Application

Summary and Analysis of RF's Mandates, Implementation, and Outcomes

Mandates

The RF program allocated funds to improve K-3 reading instruction in low-income schools by supporting the purchase of curricular materials, professional development, and additional staff support for instruction grounded in “scientifically based reading research.” Safeguards were required to ensure impartial judging of state applications and prevent federal officials from influencing specific state curricular choices. Technical assistance and program monitoring measures were required at federal, state, and district levels, and provisions were made for assessment of the program’s efficacy.

Implementation

Implementation studies concluded that RF had been implemented “as intended by the legislation” (Moss et al., 2006, p. 14; Scott, 2007b). USDOE provided guidance for states and districts in applying for grants using both federal official and contractors, and failed to effectively screen for conflicts of interest (Ashby & General Accounting Office, 2007). The guidance provided drew heavily on the NRP’s report (National Institute of Child Health and Human Development, 2000), and added several potentially extralegal requirements to RF applications, based on its view of scientifically based reading instruction. USDOE’s mismanagement of state application processes contributed to confusion among state applicants and failed to prevent USDOE officials and representatives from illegally influencing states’ curricular choices (Ashby & General Accounting Office, 2007; USDOE, Office of Inspector General, 2006). As a result of RF requirements, many schools, including those not receiving RF

grants, adopted prepackaged curriculum products, sometimes abandoning others that had comparable research backing (Manzo, 2007b).

LEAs participating in RF adopted curriculum programs under state guidance for adherence to SBRR alignment. These programs featured a 90-minute instructional block, and employed the DIBELS assessment as a diagnostic tool for student referral to intervention services (Moss et al., 2008; Scott, 2007a). RF schools provided more SBRR-aligned professional development and instructional support through reading coaches (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2008; Powell et al., 2009), and employed a variety of administrative adjustments to support implementation (Moss et al., 2006), with qualitative studies suggesting teachers and reading coaches adapted curriculum products to local conditions (Pease-Alvarez & Samway, 2008).

Outcomes

RF funds were used to purchase curriculum products, assessments, technical assistance, and support services as intended (Chambers, Lam, Mahitivanichcha, Esra, Shambaugh, Stullich, et al., 2009). SEAs and LEAs reported that RF features were important to improvements in student achievement in addition to local policies, reporting more SBRR-aligned instructional materials than their Title I counterparts (Ashby & General Accounting Office, 2007; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Scott, 2007b). More time was allotted to reading instruction in RF schools, spending more of it on the NRP's essential components of reading instruction, though these measures only displayed modest statistically significant effects on decoding measures (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006, 2008). DIBELS was widely required as a diagnostic tool for RF programs, though comparisons

to miscue analyses of children reading literature suggest it is not effective at predicting meaning-making (Arya et al., 2005; Shelton et al., 2009).

State test results and DIBELS measures showed faster growth and greater scores among RF schools than among non-RF schools, though no consistent attenuation of the achievement gap was demonstrated (Bean et al., 2010; Chapman, 2010; Dole et al., 2010; Foorman et al., 2010; Hayden et al., 2008; McGill-Franzen et al., 2006; Moss et al., 2008; Murphy et al., 2007; Skindrud & Gersten, 2006; Torgesen, 2009; Trainin et al., 2009; Trainin & Wilson, 2010; Wong-Ratcliff et al., 2010). RF schools showed no significant difference to control schools in national SAT-10 comparison (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008), though state measures demonstrated year-to-year growth (Baker et al., 2011; Chapman, 2010; Pfannenstiel et al., 2008). State ITBS scores showed no difference or lower scores than non-RF schools and year-to-year decreases on most measures (Alvermann et al., 2007; Boise State University College of Education Center for School Improvement and Policy Studies, 2007). Evaluation with the TerraNova showed no difference to non-RF schools or lower scores on most subtests (Miller et al., 2007; Westat & Learning Point Associates, 2008) Miscue analyses of RF students behavior demonstrated no significant difference to “literature based instruction” (Arya et al., 2005). RF added professional development, intervention services, and support staff for teachers, though many reported they needed more professional development to effectively serve students and core programs often did not provide much guidance in teachers efforts to adapt them to their students’ needs (Brenner & Hiebert, 2010; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006, 2008).

Interpretation of Student Testing Outcomes

State tests and DIBELS measures showed growth at faster rates and higher scores among RF schools than non-RF schools, but, due to score inflation, these scores may better represent features of RF implementation than growth in student reading skills. Comparison of RF and non-RF performance on low stakes tests is limited by the lack of an unbiased control group, but offers a less optimistic picture of RF program outcomes.

State Tests

The state tests employed in seven studies reviewed here formed the basis of “adequate yearly progress” decisions under NCLB, and teachers were under pressure to perform well on them to avoid restructurings, firings, and state takeover. With this social pressure, state tests were particularly susceptible to score inflation, where factors other than those tests are intended to measure cause higher scores (Koretz, 2008). There is evidence that RF schools’ instruction was adjusted to align better with state tests and create higher scores. In surveys, professionals in RF schools were more likely to report that their teachers were knowledgeable about SBRR instruction, and teachers were more likely to report SBRR-aligned reading materials were available and that researcher-designated SBRR-aligned instructional practices were central to their instruction (Moss et al., 2008). The NRP’s essential components of reading instruction were also well-represented in state content standards (Schenck et al., 2005), ensuring RF programs would be closely aligned, and RF programs demonstrated increased time allotted for reading instruction and greater time addressing the NRP’s five essential components (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2006, 2008; Powell et al., 2009; Scott, 2006). The close alignment of the test with the domain of RF instruction means that students’ high performance could simply represent a mastery of the narrower domain being tested rather than representing students’ mastery of a wide range of literacy skills. This seems likely when

considered in conjunction with Brenner and Hiebert's (2010) analysis that the reading material provided by third grade core reading programs was inadequate to support students skill growth, especially among poor readers. Without further investigation, it is difficult to make valid inferences about changes in students' reading skill under RF with the observed scores on state tests.

DIBELS

The widespread growth of RF students' scores on DIBELS is also likely to better represent successful implementation than literacy skill growth among students. DIBELS is designed to assess students' early literacy skills and emphasizes phonics and fluency (Li & Zhang, 2008; Roehrig et al., 2008; Shelton et al., 2009), major features of RF programs, which often included decodable readers to support phonics instruction, and which was allotted more instructional time in RF than control programs (Dudley-Marling, 2005; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Moss et al., 2008; Powell et al., 2009; Rasinski, 2006; Shelton et al., 2009). Once again, close alignment of curricular domain and the test's domain means that high or growing student scores may not represent growth of global literacy skills. The proposition that they do not represent this global growth is reinforced by the results of low-stakes tests, which present mixed results rather than the widespread gains of DIBELS. The most optimistic low-stakes scores are those of SAT-10, which demonstrated no significant difference to comparison schools in the *RF Impact Study*, and demonstrated growth for Oregon RF schools, though results were more modest than DIBELS (Baker et al., 2011; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). Evidence suggests DIBELS is predictive of students' scores on SAT-10 (e.g. Roehrig et al., 2008), but it may not be predictive of skills demonstrated or students' scores on miscue analyses of students reading literature (Arya et al., 2005; Shelton et al., 2009).

Since the telos of any reading program is students' successful interpretation of the written word, even this limited evidence should raise serious concerns about drawing inferences about global student literacy skills from DIBELS. In state comparisons using low-stakes tests, the only consistent growth in student scores was observed in measures of spelling and phonemic awareness, both closely related to the domains of RF instruction and DIBELS assessment (Alvermann et al., 2007; Miller et al., 2007). In the absence of further information, it is more defensible to argue that growth on DIBELS measures represents successful RF implementation of phonics and fluency instruction than global student skill growth.

Low-stakes Tests

As summarized above, comparisons of RF and non-RF programs using SAT-10, ITBS, and TerraNova assessments demonstrate lower scores and lower levels of student growth than state tests or DIBELS assessments. Comparisons employing SAT-10 found no change or more moderate growth than DIBELS and state test scores (Baker et al., 2011; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Pfannenstiel et al., 2008). ITBS and TerraNova results showed no difference or score decreases (Alvermann et al., 2007; Boise State University College of Education Center for School Improvement and Policy Studies, 2007; Miller et al., 2007; Westat & Learning Point Associates, 2008). Since comparisons rarely compared RF schools to schools using no RF features, they provide limited inferences about the efficacy of the RF program at increasing student achievement. These schools may have employed many features of RF that increased student performance, which were widely disseminated through state and federal professional development, and some evidence exists that there was widespread non-RF school adoption of RF core reading programs (Baker et al., 2011; Scott, 2007b). Control groups, therefore, conflate schools that employ RF practices and were provided RF levels of funding, and

schools that do not (Herlihy et al., 2009). Comparisons that can be made with this sample are also limited because RF was designed to target the highest-poverty schools, and few adequately similar schools are available for matched comparisons (e.g., Alvermann et al., 2007; B. Gamse et al., 2011). Studies that compare RF and non-RF programs are therefore susceptible to sampling and cohort biases, and should be treated with caution.

Critics (e.g., Baker et al., 2011; Herlihy et al., 2009; Mitchell et al., 2008) argue persuasively that sampling and cohort biases make it impossible to argue whether RF “works” from the *Reading First Impact Study* (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). As will be discussed below, however, there are many ways RF could have employed existing research more effectively or been left open to improvement. Appraisal of state-level evaluations using low-stakes tests to compare year-to-year growth make it difficult to argue that RF programs evidence significantly different performance than comparison schools. SAT-10, ITBS, and TerraNova scores taken together suggest stability on most scores, with consistent growth only on measures of spelling and phonics. If we grant that these programs evidence similar performance on available assessments, which is a generous interpretation of the existing evidence, we can also look to qualitative assessments of curriculum to judge the effect of RF. Evaluations suggest students spent less time engaged with print and their engagement suffered because core reading programs failed to connect to topics of interest to them (Elish-Piper & L’Allier, 2011; B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Gerstl-Pepin & Woodside-Jiron, 2005; Kersten & Pardo, 2007; Lindemer, 2005; Pease-Alvarez & Samway, 2008) Since the growth RF students demonstrated was in phonics and fluency, areas of RF focus, RF does demonstrate the possibility for instructional improvement. Considering the stability of measures

of reading comprehension, however, it is unclear that RF's resources were committed to the most efficacious areas for students.

Curricular Change Under RF

USDOE's Conception of SBRR

NRP Key Elements of Reading Instruction. The “five essential components of reading instruction” included in NCLB were drawn from the *research design* of the National Reading Panel (NRP) report (National Institute of Child Health and Human Development, 2000; Suskind, 2007), a study commissioned by Congress with the political purpose of resolving the 1980s’ “reading wars” (Cunningham, 2001). The NRP established an *a priori* framework for their review that employed primarily studies of experimental or quasi-experimental design and subdivided reading research into the five panel-developed topics that became the “essential components” of RF reading instruction (Cunningham, 2001; National Institute of Child Health and Human Development, 2000; Suskind, 2007). The NRP’s work served as the model for USDOE’s conception of SBRR under RF, just as it did for policymakers who created NCLB. The NRP’s report remains controversial among scholars of reading instruction (Cunningham, 2001; Morrow et al., 2009; Yatvin *Minority View* in National Institute of Child Health and Human Development, 2000; *The Routledge international handbook of English, language and literacy teaching*, 2010).

Criticism of the NRP Report. The NRP review has received a variety of criticisms from scholars. Joanne Yatvin, a panel member, issued a minority view in the NRP report (National Institute of Child Health and Human Development, 2000), which argued the NRP failed to address concerns such as early literacy learning and home support for literacy, and should have explored research on the influence of pre-reading literary knowledge, understanding print

conventions, and students' motivation in learning to read. Yatvin argued that the NRP should have investigated how often and when particular instruction was important to students' literacy, and that teacher-reviewers should have been employed to balance the NRP's emphasis on scientific rather than pedagogical approaches to reading instruction. Yatvin warned that "[t]opics that were never investigated will be misconstrued as failed practices" (p. 3) (National Institute of Child Health and Human Development, 2000), an idea that found expression in USDOE advice to maintain fidelity to core reading programs rather than adapt them to different student populations. Cunningham (2001) argued that the work was philosophically naïve, failed to complete a comprehensive review of the literature by ignoring qualitative evidence, failed to ask questions of when or how much phonics instruction is necessary, and overemphasized the importance of phonemic awareness instruction in the absence of evidence of its long-term benefit. Garcia & Wiese (in Morrow et al., 2009) noted that NRP failed to include research on bilingualism and bi-literacy, topics important to a nation whose schools increasingly include English language learners.

USDOE Use of the NRP Report. While USDOE's understanding of the NRP report has drawn on Kame'enui and Simmons (2000) *Consumer's Guide*, which fails to accurately relay the report's findings (Bell, 2003), USDOE adopted many features of the NRP that have drawn criticism from scholars. USDOE emphasized the importance of quantitative research in validating instruction in spite of qualitative evidence of programs' efficacy, much as NRP ignored qualitative evidence in their literature review, leading to neglect of topics best studied using qualitative means, such as social and cultural issues (Cunningham, 2001; Manzo, 2006b, 2007b; Viadero & Manzo, 2007). Appeals to the quality of research were used to justify exclusion of key programs, despite the failure of USDOE to cite research to support programs

they approved and arguments by scholars that no core reading program has been validated by research in its entirety (Brenner & Hiebert, 2010; Cunningham, 2001; Viadero & Manzo, 2007). NRP authors supported the use of decodable texts for phonics instruction without significant evidence that this less-engaging strategy for systematic phonics instruction was necessary for students' literacy (Cunningham, 2001; Elish-Piper & L'Allier, 2011; Gerstl-Pepin & Woodside-Jiron, 2005; Kersten & Pardo, 2007; Lindemer, 2005; National Institute of Child Health and Human Development, 2000; Pease-Alvarez & Samway, 2008). USDOE supported core reading programs that employed synthetic instruction with decodable readers (Cunningham, 2001; Dudley-Marling, 2005). The NRP, RF, and NCLB all failed to address many concerns about teaching English language learners (e.g., Bailey, 2010; Harper & de Jong, 2009; Morrow et al., 2009; Nesselrodt, 2007; Slavin & Cheung, 2005). Insofar as USDOE used the NRP report to make decisions about what features RF programs should employ, the NRP report led them astray. At the same time, the NRP's report supported USDOE's rhetoric of SBRR, which served as a powerful tool of legitimation. Nicholson-Goodman & Garman (2007) found that appeals of a practice being based in research served to legitimate policy mandates. Healy's (2007) belief that the five essential elements of reading instruction mandated under RF were "proven by . . . research to be necessary for a well-rounded reading education" (p. 2) shows the influence of this legitimation in the research community. Educators responding to Scott's (2007b) survey often emphasized fidelity to research-based curriculum products as important to the best student outcomes.

The Research Base of RF Curricula. The core reading programs adopted under RF were the subject of controversy. The emphasis on quantitative studies in SBRR requirements attenuated the research base that professionals could formally draw on, and created confusion as

states made RF applications (Cunningham, 2001; Foorman & Nixon, 2006). Kame'enui and Simmons' (2000) *Consumer's Guide* was adopted by many states and emphasizes the use of decodable readers and their research support, despite NRP's failure to find sufficient evidence to endorse them (Bell, 2003; Cunningham, 2001; National Institute of Child Health and Human Development, 2000). Scholars Brenner and Hiebert (2010) argued that no core reading program was entirely validated by research, and found that the curricula approved under RF lacked sufficient opportunity for students to read. Three of the most often employed core reading programs under RF—McGraw-Hill's Open Court, Scott Foresman Reading, and Houghton Mifflin Reading—were not supported by later reviews due to a dearth of research to validate them (Institute of Education Sciences (ED), 2007). Certain programs, including *Success for All*, *Reading Recovery*, *Kaplan SpellRead*, and *Start Making a Reader Today* were barred from or employed much less among RF programs than among schools generally, despite being rated by independent organizations as research-based (Healy, 2007; Institute of Education Sciences (ED), 2007; Manzo, 2007b, 2007c). USDOE emphasis on fidelity in executing core reading programs (Dudley-Marling, 2005) was found to be confusing for teachers to implement due to a lack of guidance in provided materials (Brenner & Hiebert, 2010), and USDOE professional development disregarded NRP advice that teachers adapt instruction to students' individual needs (National Institute of Child Health and Human Development, 2000). USDOE's approved curricula was therefore often constructed of less-than research-validated claims and programs, and the advice it provided ran against that of researchers and the creators of core reading programs (Pease-Alvarez & Samway, 2008).

RF teachers in Moss and colleagues' (2006) survey of RF implementation were modestly more likely to rate instructional practices that researchers felt reflected SBRR as central to their

reading instruction. The non-SBRR practices that researchers identified at the kindergarten level included teachers reading stories aloud to the whole class, students using the dictionary to find word meanings, and teaching phonemic awareness skills in the course of reading (Moss et al., 2006). These activities, SBRR or not, can have important functions for teachers, whose goal may not simply be to increase test scores, but to foster literate habits among students.

Impact of the RF Program

Instructional Time

RF increased the daily time allotted to reading programs to and beyond the 90-minute instructional block mandated by USDOE, and demonstrated greater observed instructional time in the NRP's five essential components of reading instruction, explicit instructions, and practice in the five essential components than comparison schools (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). This increase would be admirable if it necessarily indicated improvements in the teaching and learning of ELA, but scholars warn that it does not reflect the quality of educational activities (Brenner & Hiebert, 2010; Corey et al., 2012). Students' scores on SAT-10 suggested a 0.07 point increase in scale scores for a one-minute increase in average time devoted to the daily reading block, amounting to a modest 1.3 point increase each year based on the additional 18.5 minutes more reported by RF schools (B. C. Gamse, Jacob, Horst, Boulay, & Unlu, 2008). The widespread growth in RF students' scores on DIBELS and moderate gains on low-stakes phonics measures suggest that the additional time allotted to reading instruction under RF did contribute to growth in students' phonics skills, in line with USDOE's desired curriculum. Unfortunately, that focus on phonics has not translated into higher reading comprehension scores on low-stakes tests than comparison schools. This may be because, as researchers have argued (Cunningham, 2001; Yatvin in National Institute of Child Health and Human Development,

2000), they failed to ask questions of when and how much phonics instruction was appropriate for students, and failed to implement it in creative, engaging ways. Studies suggest that extending phonics instruction into higher grades may be a poor way to improve student literacy skills (e.g., Edmonds et al., 2009), yet RF programs seem to have done just that.

In addition to the questionable efficacy of RF's emphasis on phonics, the time it mustered for reading instruction was likely taken from science, the humanities, and arts electives, as other schools demonstrated under NCLB (McMurrer, 2008). This evidence suggests that students were provided few opportunities to learn science, humanities, and arts background knowledge, which has been found to be a better predictor of text comprehension than reading skills (Willingham, 2009). Changes in instruction may have generated some growth in students' phonics and spelling skills, but the costs—in stable reading comprehension scores, in attenuation of means teachers can employ to engage students, in students' engagement with print, and in background knowledge they may have learned—were too great.

The Expansion of Administrative Support

The RF program encouraged various beneficial expansions of administrative support. Teachers in RF schools reported receiving 12.1 more hours of professional development than expected in non-RF schools; covered more of the NRP's five essential components of reading instruction (B. C. Gamse, Jacob, Horst, Boulay, & Unlu, 2008); reported more time to use assessment data to plan instruction; to observe reading instruction in other classrooms; to collaborate on lesson planning and instruction; and to receive instruction from a reading coach (Moss et al., 2006).

Reading coaches were a powerful means for RF implementation and instructional support by supporting teachers' use of SBRR instructional materials, providing interventions, providing

professional development, and supporting teachers' use of student achievement data to inform instruction (Moss et al., 2006, 2008; Scott, 2007a). Teachers reported they were more likely to adopt instructional changes suggested by reading coaches than those from other sources (Coburn & Woulfin, 2012). RF reading coaches also reported spending more time in that role than Title I coaches (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008). Despite these additional administrative supports under RF, 80% of teachers reported they needed additional professional development to provide effective interventions for struggling readers, a figure identical to that of Title I comparison schools (Moss et al., 2006).

The Use of Core Reading Programs

One of the most overt signals of RF alignment was the use of RF-required core reading programs. USDOE visions of what constituted SBRR, though flawed, attenuated the curriculum products employed by discouraging use of certain programs (Healy, 2007; Manzo, 2007b; Viadero & Manzo, 2007). The emphasis on grade-level and whole-class instruction under approved curricula may have interfered with the preparation for reading and opportunity for independent practice that would allow struggling readers to improve (McGill-Franzen et al., 2006). Approved programs often employed decodable readers per *Consumer's Guide* and NRP suggestion, but critics suggest that this was a less-engaging means to teach phonics than teachers might have employed (Cunningham, 2001). Research further suggests that teachers may not have felt prepared to implement core programs because programs provided limited guidance on their implementation or adaptations to local conditions.

Rather than simply maintain fidelity to core reading programs as suggested by USDOE guidance (Brenner & Hiebert, 2010; USDOE, Office of Elementary and Secondary Education, 2002), qualitative studies have found that teachers and literacy coaches adapted programs to

local visions of best practice or students' individual needs, sometimes at the prompting of core reading program technical support (Elish-Piper & L'Allier, 2011; Gerstl-Pepin & Woodside-Jiron, 2005; Kersten & Pardo, 2007; Pease-Alvarez & Samway, 2008). Teachers among lower-income schools reported greater pressure to maintain fidelity due to NCLB AYP provisions (McCartney, 2008). Teacher adaptations garnered support from administrators skeptical of RF programs, and conflict from those who supported them (Pease-Alvarez & Samway, 2008). In many cases, adaptation of core reading programs occurred when educators identified programs' failure to engage or meet the needs of diverse students (Elish-Piper & L'Allier, 2011; Gerstl-Pepin & Woodside-Jiron, 2005; Kersten & Pardo, 2007; Lindemer, 2005; Pease-Alvarez & Samway, 2008)

RF teachers were more likely to report the use of assessments for grouping students during instruction, identifying students who need interventions, and measuring student progress were helpful in improving student achievement (Moss et al., 2008). In order to do this, most RF schools employed DIBELS, as recommended by USDOE (Scott & Fagan, 2005), with its emphasis on phonics (Li & Zhang, 2008). Researchers, however, have argued that DIBELS provides limited diagnostic information to teachers compared to other tests, fails to take into account the impact of social and cultural factors on reading development, and over-emphasizes speed in reading rather than comprehension (Li & Zhang, 2008; Morrow et al., 2009). The USDOE (2006) *NCLB Executive Summary* suggests that the RTI framework adopted under RF will create "reduced identification of children for special education services due to a lack of appropriate reading instruction in their early years" (p. 3). This seems to have been accurate, as fewer students were identified as learning disabled under RF in Florida than were before RF implementation (Torgesen, 2009). This may also have had negative consequences. Title I schools

were more likely to report having certified special education teachers provide recommendations to plan instruction for struggling readers, and RF teachers in K-2 were less likely to report time being set aside to coordinate reading instruction for special education students (Moss et al., 2006). In any case, RF had the benefit of getting more struggling readers into intervention services than comparison schools and increased use of assessment to evaluate readers' difficulties (Moss et al., 2008).

Other ELA Concerns

The RF program failed to expand writing instruction in early grades curriculum, which had been growing before NCLB (Squire in Flood, International Reading Association, & National Council of Teachers of English, 2003). This instruction would have laid the groundwork for students' future writing instruction, which is important to students' college experiences and future careers (Lapp, Flood, & Farnan, 2008). From 2002 to 2007, 8th grade student reports of writing for school gathered by the NAEP showed a small but significant drop (Applebee & Langer, 2009). This cannot be attributed to RF directly, but RF influence on upper-grades instruction may have contributed, and NAEP drops highlight lost opportunities to expand students' reading skills.

Valli, Croninger, and Buese's (2012) qualitative analysis of teaching change in school years 2002-3, 2003-4, and 2004-5 found a trend of decreasing cognitive demand and increasing focus on basic skills as NCLB policies were implemented, although they do not distinguish effects of high-stakes testing measures and RF grant effects. While the highly explicit instruction advocated under RF may allow easier acquisition of basic concepts, teachers should provide students opportunities to engage with difficult cognitive tasks to expand their skills and abilities. This can prompt students to connect classroom concepts with practical uses outside the

classroom, engage their drives for mastery and autonomy, and support independent practice of skills learned in school (Ryan & Deci, 2000).

RF as Public Policy

Successful Implementation

RF's implementation was successful, reaching 1,809 districts and 5,880 schools in all 50 states and all U.S.-affiliated entities directly, 67% of whom reported having to change reading instruction to meet RF requirements (B. C. Gamse, Jacob, Horst, Boulay, Unlu, et al., 2008; Scott, 2007b). Many other LEAs are likely to have adopted RF programs or features (Baker et al., 2011) as a result of Title I coordination, participation in RF professional development, AYP pressures under NCLB, and extension of RF features to upper grades (Scott, 2007b). The use of rhetoric that emphasized RF's basis in SBRR (e.g., USDOE, Office of Elementary and Secondary Education, 2002; USDOE, 2002) served to support implementation among administrators and educators, for whom the nominal research basis functioned to legitimate mandates (Nicholson–Goodman & Garman, 2007). This rhetoric was supported by the use of DIBELS as an assessment of RF schools, which demonstrated widespread score growth and likely contributed to positive educator reviews of RF programs' effectiveness (Scott, 2007b). RF's stringent requirements were supported by \$5.39 billion of professional development and technical assistance that enabled successful implementation. USDOE and SEAs provided technical assistance and professional development, while LEAs employed core programs, assessments, intervention services, and reading assessments to support students' skill growth.

Ethical Failure

RF's successful implementation is obscured by its ethical failure. High levels of executive influence on the shape of RF programs and failures to screen for conflicts of interest

allowed USDOE officials and their representatives the opportunity to enforce their visions of appropriate reading instruction, which have had small positive effects for drastically high costs. The lack of oversight of USDOE in its treatment of state RF applications contributed to confusion among states over program requirements and contributed to the erosion of public trust that precipitated RF's drastic budget cuts and cancellation. RF had some standardizing effect on low-income schools, with many adopting similar core reading programs and assessments. RF therefore prevented the use of states as laboratories for policy experiments that might have occurred under traditional American federalism (Healy, 2007).

Chapter 4: Conclusions & Recommendations

Summary of Conclusions

Mandates, Implementation, and Outcomes

RF was established as part of NCLB to serve as a model of how successful reading instruction would allow low-income students to achieve at similar levels to more advantaged students. It promoted the adoption of selected core reading programs, assessments, and intervention services. These tools created a greater phonics emphasis in K-3 reading instruction, which resulted in growth on phonics measures but stability compared to comparison schools on most low-stakes measures of reading comprehension. The time allotted for RF's phonics-heavy reading instruction likely detracted from student learning in key areas of background knowledge, such as science, the humanities, and the arts, in addition to ignoring teachers' ongoing concerns about early grades writing instruction. Qualitative studies suggest that educators adapted core reading programs to local circumstances in spite of USDOE emphasis on maintaining fidelity, which likely prevented core reading programs from alienating students entirely, though engagement likely suffered. Struggling readers were swiftly provided support under the RTI framework and intervention services provided by RF grants, which likely allowed educators to quickly address students' weaknesses identified by testing. The testing products employed, however, are subject to score inflation and over-represent the phonics in RF curriculum enough that they are dubious evaluations of students' overall skill growth.

Interpretation of Outcomes

The general stability of students' scores on low stakes tests measuring literacy skills other than phonics suggests that RF's curriculum shifts moved the emphasis of lower-grades reading instruction to phonics rather than comprehension. If we interpret scores optimistically as a sign

that students' skill growth in RF and comparison schools was similar, there are signs of unmeasured negative effects of students' learning. Lower student engagement in qualitative studies suggests that instruction under RF may not have been as successful in connecting to students' motivations or drives for autonomy and mastery. The adjustment of core reading programs by local teachers and reading coaches suggests that true fidelity to core reading programs would have elicited worse effects on students' motivation. Researchers' criticisms of USDOE's conception of SBRR identify where improved understanding of what *is* research-based would have helped to create core reading programs that addressed phonics concerns and those of reading comprehension, providing the opportunity for wider growth of students' skills.

RF as Policy

RF did not fulfill its mission of bringing low-income students to similar growth as their more-advantaged peers. It remains, however, a successfully-implemented and -supported policy. Measures of instructional time, students' standardized test scores, and qualitative reports suggest widespread adherence to core reading programs, assessments, and interventions, with support from federal and state agencies in professional development and technical assistance. The use of SBRR rhetoric throughout the law and the support of literacy coaches supported local implementation, and the use of DIBELS assessments helped to convince local educators of improvement under the program, regardless of what can be validly interpreted from them.

RF failed notably, however, on ethical grounds. The failure to institute adequate guards for conflicts of interest provided opportunities for USDOE officials and contractors to project their view of SBRR instruction onto the national stage, and mismanagement of the state application process contributed to early confusion among SEAs.

Recommendations

Further Study

Further study should investigate what features of RF programs contributed to students' score growth on low-stakes tests, how well DIBELS scores correlate with phonics instruction, and whether less time-intensive methods than decodable readers are sufficient to meet students' phonics-learning needs so that curricula can be adjusted for more growth in reading comprehension and other indicators of student skills. Many questions about appropriate early-grades instruction remain, and further study would allow better understanding of how phonics can be balanced with other important literacy concerns.

For Teachers

Teachers should continue to adapt reading programs to the needs of their students. Contextual factors may affect the efficacy of reading programs, and teachers who work in conjunction with the advice of reading coaches, administrators, and research can create instruction that is more appropriate to their pedagogical situations.

In the event that teachers face the powerful legitimization of SBRR rhetoric in their own districts, they can use research that validates their desired practices to help convince other professionals. Qualitative studies suggest that core reading program technical assistance and literacy coaches can be important factors in curriculum adaptation in spite of SBRR rhetoric.

For Policymakers

Policymakers should make themselves aware of the political ties between researchers and educational service companies. They should maintain skepticism in the face of SBRR rhetoric, and ensure adequate oversight of executive agencies. Less prescriptive policies may provide executive agencies fewer opportunities for overreach, and the wider research community remains an important resource for further information about effective instruction and oversight.

Importance of This Work

This work contributes to educational knowledge by describing and analyzing the changes in the ELA curriculum as a result of the RF program. It provides insight into the effects of RF's prescriptive requirements on reading instruction and their outcomes. It traces the changes that occurred as RF was translated from federal policy to classroom practice, and analyzes their impact with important implications for researchers, teachers, and policymakers.

References

- Alvermann, D., Commeyras, M., Cramer, S., Harnish, D., Pollack, S., & Alagoz, C. (2007, August 15). Evaluation of Georgia Reading First Implementation, Progress and Impact 2006-07 Year Three (Cohort 1) Year One (Cohort 2). USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/performance.html>
- Applebee, A. N., & Langer, J. A. (2009). EJ Extra: What Is Happening in the Teaching of Writing? *The English Journal*, 98(5), 18–28. doi:10.2307/40503291
- Arya, P., Martens, P., Wilson, P. G., Altwerger, B., Jin, L., Laster, B., & Lang, D. (2005). Reclaiming Literacy Instruction: Evidence in Support of Literature-Based Programs. *Language Arts*, 83(1), 63–72.
- Ashby, C. M., & General Accounting Office, W. (2007). *Reading First: States Report Improvements in Reading Instruction, but Additional Procedures Would Clarify Education's Role in Ensuring Proper Implementation by States. Report to Congressional Requesters. GAO-07-161*. Government Accountability Office.
- Bailey, L. (2010). The Impact of Sustained, Standards-Based Professional Learning on Second and Third Grade Teachers' Content and Pedagogical Knowledge in Integrated Mathematics. *Early Childhood Education Journal*, 38(2), 123–132. doi:10.1007/s10643-010-0389-x
- Baker, S. K., Smolkowski, K., Smith, J. M., Fien, H., Kame'enui, E. J., & Beck, C. T. (2011). The Impact of Oregon Reading First on Student Reading Outcomes. *Elementary School Journal*, 112(2), 307–331.

- Bean, R., Draper, J., Turner, G., & Zigmond, N. (2010). Reading First in Pennsylvania: Achievement Findings after Five Years. *Journal of Literacy Research*, 42(1), 5–26.
doi:10.1080/10862960903583244
- Bell, M. (2003). The International Reading Association's review of Reading First grant recipients. *Reading Teacher*, 56(7), 670.
- Boise State University College of Education Center for School Improvement and Policy Studies. (2007). Idaho Reading First 2006-2007 Evaluation. USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/performance.html>
- Bonds, D. E. (2010, January 1). *Exploring Reading First Program Implementation across Colorado Schools with Differing Achievement Results*. ProQuest LLC.
- Bowers, M. M. (2011). *A Study of Reading First Implementation and Literacy Performance of Students in Kindergarten through Fifth Grade in Lancaster School District*. ProQuest LLC.
- Brenner, D., & Hiebert, E. H. (2010). If I Follow the Teachers' Editions, Isn't That Enough? Analyzing Reading Volume in Six Core Reading Programs. *Elementary School Journal*, 110(3), 347–363.
- Brown-Wright, L., Tyler, K. M., Stevens-Watkins, D., Thomas, D., Mulder, S., Hughes, T., ... Smith, L. T. (2013). Investigating the Link Between Home–School Dissonance and Academic Cheating Among High School Students. *Urban Education*, 48(2), 314–334.
doi:10.1177/0042085912469203
- Callow-Heusser, C., & Chapman, H. (2007, November 30). Bureau of Indian Education's (IE) Reading First Program: 2006-07 Summative Evaluation Report. USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/evaluationreports/bia.pdf>

- Chambers, J. G., Lam, I., Mahitivanichcha, K., Esra, P., Shambaugh, L., & Stullich, S. (2009). State and Local Implementation of the No Child Left Behind Act. Volume VI--Targeting and Uses of Federal Education Funds. *US Department of Education*. Retrieved from <http://eric.ed.gov/?id=ED504207>
- Chambers, J. G., Lam, I., Mahitivanichcha, K., Esra, P., Shambaugh, L., Stullich, S., & Office of Planning, E. and P. D. (ED). (2009). *State and Local Implementation of the No Child Left Behind Act. Volume VI--Targeting and Uses of Federal Education Funds*. US Department of Education.
- Chapman, H. J. (2010, January 1). *Factors Affecting Reading Outcomes across Time in Bureau of Indian Education Reading First Schools*. ProQuest LLC.
- Coburn, C. E., & Woulfin, S. L. (2012). Reading Coaches and the Relationship Between Policy and Practice. *Reading Research Quarterly*, 47(1), 5–30. doi:10.1002/RRQ.008
- Corey, D. L., Phelps, G., Ball, D. L., Demonte, J., & Harrison, D. (2012). Explaining Variation in Instructional Time An Application of Quantile Regression. *Educational Evaluation and Policy Analysis*, 34(2), 146–163. doi:10.3102/0162373711431098
- CTB/McGraw-Hill LLC. (2013). TerraNova 3: The new standard in achievement. McGraw-Hill Education. Retrieved from <http://www.ctb.com/ctb.com/control/productFamilyViewAction?productFamilyId=449&p=products>
- Cunningham, J. W. (2001). Report of the National Reading Panel: teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (Book Review) (Undetermined). *Reading Research Quarterly*, 36(3), 326–335.

- Day-Meeks, A. L. (2011, January 1). *Assessing the Impact of Reading First Programs on Student Achievement in K-3 Classrooms in Selected Mississippi Schools*. ProQuest LLC.
- Dole, J. A., Hosp, J. L., Nelson, K. L., & Hosp, M. K. (2010). Second Opinions on the Reading First Initiative: The View from Utah. *Journal of Literacy Research*, 42(1), 27–48.
doi:10.1080/10862960903583277
- Dudley-Marling, C. (2005). Disrespecting Teachers: Troubling Developments in Reading Instruction. *English Education*, 37(4), 272–279. doi:10.2307/40173200
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A Synthesis of Reading Interventions and Effects on Reading Comprehension Outcomes for Older Struggling Readers. *Review of Educational Research*, 79(1), 262–300. doi:10.2307/40071166
- Elias, E. I. (2009, January 1). *The Lived Experiences of Six First-Grade Teachers Using Reading Mastery Plus Curriculum in High Poverty Schools*. ProQuest LLC.
- Elish-Piper, L., & L’Allier, S. K. (2011). Examining the Relationship Between Literacy Coaching and Student Reading Gains in Grades K-3. *Elementary School Journal*, 112(1), 83–106.
- Elmore, R. F. (2004). *School reform from the inside out: policy, practice, and performance*. Cambridge, Mass: Harvard Education Press.
- Flood, J., International Reading Association, & National Council of Teachers of English. (2003). *Handbook of research on teaching the English language arts*. Mahwah, N.J.: L. Erlbaum Associates.
- Foorman, B. R., & Nixon, S. M. (2006). The Influence of Public Policy on Reading Research and Practice. *Topics in Language Disorders*, 26(2), 157–171.

- Foorman, B. R., Petscher, Y., Lefsky, E. B., & Toste, J. R. (2010). Reading First in Florida: Five Years of Improvement. *Journal of Literacy Research*, 42(1), 71–93.
doi:10.1080/10862960903583202
- Gamse, B., Boulay, B., Rulf-Fountain, A., Unlu, F., & Society for Research on Educational Effectiveness (SREE). (2011). *Findings from the Reading First Implementation Study, 2008-09*. Society for Research on Educational Effectiveness.
- Gamse, B. C., Bloom, H. S., Kemple, J. J., Jacob, R. T., & Institute of Education Sciences (ED), W. (2008a). *Reading First Impact Study: Interim Report. Executive Summary. NCEE 2008-4019*. National Center for Education Evaluation and Regional Assistance.
- Gamse, B. C., Bloom, H. S., Kemple, J. J., Jacob, R. T., & Institute of Education Sciences (ED), W. (2008b). *Reading First Impact Study: Interim Report. NCEE 2008-4016*. National Center for Education Evaluation and Regional Assistance.
- Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). Reading First Impact Study Executive Summary (NCEE 2009-4039). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., Unlu, F., & Institute of Education Sciences (ED), N. C. for E. E. and R. A. (2008). *Reading First Impact Study. Final Report. NCEE 2009-4038*. National Center for Education Evaluation and Regional Assistance.
- Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher Study Group: Impact of the Professional Development Model on Reading Instruction and Student Outcomes in First Grade Classrooms. *American Educational Research Journal*, 47(3), 694–739.

- Gerstl-Pepin, C. I., & Woodside-Jiron, H. (2005). Tensions between the “Science” of Reading and a “Love of Learning”: One High-Poverty School’s Struggle with NCLB. *Equity and Excellence in Education*, 38(3), 232–241.
- Glenn, D., & Brainard, J. (2007). House Panel Cuts Reading-First Funds. *Chronicle of Higher Education*, 53(46), A21–A21.
- Grunwald, M. (2006, October 1). Billions for an Inside Game on Reading. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/wp-dyn/content/article/2006/09/29/AR2006092901333.html>
- Harper, C. A., & de Jong, E. J. (2009). English language teacher expertise: the elephant in the room. *Language & Education: An International Journal*, 23(2), 137–151.
doi:10.1080/09500780802152788
- Hayden, E., Trainin, G., Javorsky, K., Murphy-Yagil, M., & Cook, K. (2008). *Nebraska Reading First 2007-08 Annual Report*.
- Healy, K. C. (2007). Reading First, Federalism Second? How a Billion Dollar NCLB Program Disrupts Federalism. *Columbia Journal of Law & Social Problems*, 41(2), 147–175.
- Herlihy, C., Kemple, J. J., Bloom, H. S., Zhu, P., & Berlin, G. (2009). Understanding Reading First: What we know, what we don’t, and what’s next. MDRC. Retrieved from http://www.mdrc.org/sites/default/files/understanding_reading_first.pdf
- Hewitt, D. T. (2011). Reauthorize, Revise, and Remember: Refocusing the No Child Left Behind Act To Fulfill Brown’s Promise. *Yale Law & Policy Review*, 30(1), 169–194.
- Houghton Mifflin Harcourt. (2010). Iowa Tests of Basic Skills (ITBS) Forms A, B, and C. The Riverside Publishing Company. Retrieved from <http://www.riverpub.com/products/itbs/>

- Institute of Education Sciences (ED). (2007, August 13). What Works Clearinghouse Topic Report: Beginning Reading. USDOE. Retrieved from <http://education.ucf.edu/mirc/Research/IES%20-%20Beginning%20Reading.pdf>
- Jeynes, W. (2007). *American educational history: school, society, and the common good*. Thousand Oaks: SAGE Publications.
- Kame'enui, E. J., & Simmons, D. C. (2000). Consumer's guide to evaluating a core reading program: a critical elements analysis. University of Oregon, Institute for the Development of Educational Achievement. Retrieved from http://reading.uoregon.edu/cia/curricula/con_guide.php
- Kersten, J., & Pardo, L. (2007). Finessing and Hybridizing: Innovative Literacy Practices in Reading First Classrooms. *Reading Teacher*, 61(2), 146–154.
- Koretz, D. M. (2008). *Measuring up what educational testing really tells us*. Cambridge, Mass.: Harvard University Press. Retrieved from <http://site.ebrary.com/id/10313837>
- Lapp, D., Flood, J., & Farnan, N. (2008). *Content area reading and learning: instructional strategies*. New York: Lawrence Erlbaum.
- Li, X., & Zhang, M. (2008). Reconciling DIBELS and OSELA: what every childhood educator should know. *Journal of Research in Childhood Education*, 23(1), 41+.
- Lindemer, C. L. (2005). Pulling It Together: Meeting the Needs of Struggling Readers in a Standard's Based Title I Reading Program. *Ohio Reading Teacher*, 37(1), 37–47.
- Manzo, K. K. (2002). Majority of States Told to Revise Reading Plans. *Education Week*, 22(5), 10.
- Manzo, K. K. (2004). Reading Programs Bear Similarities Across the States. *Education Week*, 23(21), 1–13.

- Manzo, K. K. (2005). States Pressed To Refashion Reading First Grant Designs. *Education Week*, 25(2), 1–25.
- Manzo, K. K. (2006a). Audit Faults Wisconsin’s “Reading First” Grant Process. *Education Week*, 26(10), 21–21.
- Manzo, K. K. (2006b). Experts, Policymakers at Odds Over Reading. *Education Week*, 25(21), 8–8.
- Manzo, K. K. (2007a). E-Mails Reveal Federal Reach Over Reading. *Education Week*, 26(24), 1–19.
- Manzo, K. K. (2007b). Federal Reading Review Overlooks Popular Texts. *Education Week*, 27(1), 8–8.
- Manzo, K. K. (2007c). Reading Probe Will Continue On Capitol Hill. *Education Week*, 26(31), 1–17.
- Manzo, K. K. (2007d). Reading Results Hard to Translate, Panel Concludes. *Education Week*, 27(1), 9–9.
- McCallion, G. (2008, February 6). Congressional Research Report RL33246: Reading First: Implementation Issues and Controversies. WikiLeaks. Retrieved from <http://crs.wikileaks-press.org/RL33246.pdf>
- McCartney, S. J. (2008). The Impact of No Child Left Behind on Teachers’ Writing Instruction. *Written Communication*, 25(4), 462–505. doi:10.1177/0741088308322554
- McDermott, K. A. (2011). *High-stakes reform: the politics of educational accountability*. Washington, DC: Georgetown University Press.

- McGill-Franzen, A., Zmach, C., Solic, K., & Zeig, J. L. (2006). The Confluence of Two Policy Mandates: Core Reading Programs and Third-Grade Retention in Florida. *Elementary School Journal*, 107(1), 67–93.
- McGuinn, P. J. (2006). *No Child Left Behind and the transformation of federal education policy, 1965-2005*. Lawrence, Kan: University Press of Kansas.
- McMurrer, J. (2008). Instructional Time in Elementary Schools: A Closer Look at Changes for Specific Subjects. *From the Capital to the Classroom: Year 5 of the No Child Left Behind Act*.
- Miller, S., Swanlund, A., Butler, J., Cisneros, A., Davidson-Gibbs, D., & Lessem, S. (2007, November). Wisconsin Reading First Year 4 (2006-2007). USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/performance.html>
- Mitchell, K., Brady, S., Cutting, L., Carlo, M., Gutierrez, K., Lesaux, N. K., ... Vellutino, F. (2008, August 20). Response to the Reading First Impact Study Interim Report. USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/statement.pdf>
- Morrow, L. M., Rueda, R., & Lapp, D. (2009). *Handbook of research on literacy and diversity*. New York: Guilford Press.
- Moss, M., Fountain, A. R., Boulay, B., Horst, M., Rodger, C., Brown-Lyons, M., & Office of Planning, E. and P. D. (ED). (2008). *Reading First Implementation Evaluation: Final Report*. US Department of Education.
- Moss, M., Jacob, R., Boulay, B., Horst, M., Poulos, J., Department of Education, W., & Abt Associates, I. (2006). *Reading First Implementation Evaluation: Interim Report*. US Department of Education.

- Murphy, M., Trainin, G., Yagil, O., Javorsky, K., & Hayden, E. (2007). *Nebraska Reading First: Three Years of Implementation, 2006-2007. Annual Report*. Online Submission.
- National Center for Education Statistics. (2012, September). NAEP overview. Institute of Education Sciences. Retrieved from <http://nces.ed.gov/nationsreportcard/about/>
- National Center on Response to Intervention. (n.d.). What is RTI? National Center on Response to Intervention. Retrieved from National Center on Response to Intervention
- National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel: Teaching Children to Read: An evidenc-based assessmentof the scientific research literature on reading and its implications for reading instruction. National Institutes of Health.
- Nelsestuen, K., Scott, C., Hanita, M., Robinson, L., Coskie, T., & Regional Educational Laboratory at Education Northwest (ED). (2009). *New and Experienced Teachers in a School Reform Initiative: The Example of Reading First. Issues & Answers. REL 2009-082*. Regional Educational Laboratory at Education Northwest.
- Nesselrodt, P. S. (2007). Ramping Up to Meet NCLB Mandates by Creating an ESL Program Reflecting Effective Schools Research. *Journal of Education for Students Placed at Risk (JESPAR)*, 12(4), 441–457. doi:10.1080/10824660701759049
- Nicholson–Goodman, J., & Garman, N. B. (2007). Mapping practitioner perceptions of “It”’s research based’: scientific discourse, speech acts and the use and abuse of research. *International Journal of Leadership in Education*, 10(3), 283–299. doi:10.1080/13603120701257297
- No Child Left Behind Act of 2001 (Public Law 107-110. (2002). USDOE. Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf>

- No child left behind?: the politics and practice of school accountability*. (2003). Washington, D.C: Brookings Institution Press.
- Paige, R. (2003, October 27). Real Education Reform. *New York Times*. New York. Retrieved from <http://www.nytimes.com/2003/10/27/opinion/1-real-education-reform-839027.html>
- Pearson Education. (2013a). Group Reading Assessment and Diagnostic Evaluation (GRADE). Pearson. Retrieved from <http://www.pearsonassessments.com/learningassessments/products/100000646/group-reading-assessment-and-diagnostic-evaluation-grade-grade.html?Pid=GRADE>
- Pearson Education. (2013b). Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4). Pearson. Retrieved from <http://www.pearsonclinical.com/language/products/100000501/peabody-picture-vocabulary-test-fourth-edition-ppvt4.html?Pid=PAa30700>
- Pease-Alvarez, L., & Samway, K. D. (2008). Negotiating a Top-Down Reading Program Mandate: The Experiences of One School. *Language Arts*, 86(1), 32–41.
- Pfannenstiel, J., Seltzer, D., & Yarnell, V. (2008). Evaluation of Kansas Reading First Grant: Three-Year Report of Achievement Results. USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/evaluationreports/index.html>
- Powell, D., Higgins, H. J., Aram, R., & Freed, A. (2009). Impact of No Child Left Behind on Curriculum and Instruction in Rural Schools. *Rural Educator*, 31(1), 19–28.
- Rasinski, T. (2006). Reading Fluency Instruction: Moving Beyond Accuracy, Automaticity, and Prosody. *The Reading Teacher*, 59(7), 704–706. doi:10.1598/RT.59.7.10
- Robelen, E. W. (2002). An ESEA Primer. *Education Week*, 21(16), 28.

- Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS Oral Reading Fluency Measure for Predicting Third Grade Reading Comprehension Outcomes. *Journal of School Psychology, 46*(3), 343–366.
doi:10.1016/j.jsp.2007.06.006
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology, 25*(1), 54–67.
doi:10.1006/ceps.1999.1020
- Schenck, E. A., Walker, D. R., Nagel, C. R., Webb, L. C., RMC Research Corp., A., McKenzie Group, W., & Department of Education, W. (2005). *Analysis of State K-3 Reading Standards and Assessments. Final Report.* US Department of Education.
- Schilling, S. G., Carlisle, J. F., Scott, S. E., & Ji Zeng. (2007). Are Fluency Measures Accurate Predictors of Reading Achievement? *Elementary School Journal, 107*(5), 429–448.
- Scott, C. (2006). Keeping Watch on Reading First. *Center on Education Policy, 1–19.*
- Scott, C. (2007a). Moving Beyond Identification: Assisting Schools in Improvement. *From the Capital to the Classroom: Year 5 of the No Child Left Behind Act, 1–37.*
- Scott, C. (2007b). NCLB Year 5: Reading First: Locally Appreciated, Nationally Troubled. *Center on Education Policy.* Retrieved from <http://www.cepdc.org/displayDocument.cfm?DocumentID=310>
- Scott, C., & Fagan, T. (2005). Ensuring Academic Rigor or Inducing Rigor Mortis?: Issues to Watch in Reading First. *Center on Education Policy, 1–16.*
- Shelton, N. R., Altwerger, B., & Jordan, N. (2009). Does DIBELS Put Reading First? *Literacy Research and Instruction, 48*(2), 137–148.

- Skindrud, K., & Gersten, R. (2006). An Evaluation of Two Contrasting Approaches for Improving Reading Achievement in a Large Urban District. *Elementary School Journal*, 106(5), 389–407.
- Slavin, R. E., & Cheung, A. (2005). A Synthesis of Research on Language of Reading Instruction for English Language Learners. *Review of Educational Research*, 75(2), 247–284. doi:10.2307/3516050
- Stewart, M. T. (2004). Early literacy instruction in the climate of No Child Left Behind. *Reading Teacher*, 57(8), 732–743.
- Suskind, D. C. (2007). Going Public: NCLB and Literacy Practices in Teacher Education. *Language Arts*, 84(5), 450–455.
- The Routledge international handbook of English, language and literacy teaching*. (2010). London ; New York: Routledge.
- Torgesen, J. K. (2009). The Response to Intervention Instructional Model: Some Outcomes From a Large-Scale Implementation in Reading First Schools. *Child Development Perspectives*, 3(1), 38–40. doi:10.1111/j.1750-8606.2009.00073.x
- Trainin, G., Javorsky, K., Murphy, M., & Wilson, K. (2009). *Nebraska Reading First: Year Five of Implementation--2008-2009. Annual Report*. Online Submission.
- Trainin, G., & Wilson, K. (2010). *Nebraska Reading First: Year Six of Implementation--2009-2010. Annual Report*. Online Submission.
- USDOE. (2002). No Child Left Behind: Elementary and Secondary Education Act (ESEA). USDOE. Retrieved from <http://www2.ed.gov/nclb/landing.jhtml>
- USDOE. (2006). NCLB Executive Summary. Retrieved from <http://www.ed.gov/nclb/overview/intro/execsumm.pdf>

- USDOE. (2009). Reading First. USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/index.html>
- USDOE. (2011, August 31). Improving basic programs operated by local educational agencies (title I, part A). USDOE. Retrieved from <http://www2.ed.gov/programs/titleiparta/index.html>
- USDOE, Office of Elementary and Secondary Education. (2002). Guidance for the Reading First Program. Retrieved from <http://www2.ed.gov/programs/readingfirst/guidance.pdf>
- USDOE, Office of Inspector General. (2006, September). The Reading First Program's Grant Application Process: Final inspection Report. USDOE. Retrieved from <http://www2.ed.gov/about/offices/list/oig/aireports/i13f0017.pdf>
- USDOE, Office of Inspector General. (2007a, February). The Department's Administration of Selected Aspects of the Reading First Program: Final Audit Report. USDOE. Retrieved from <http://www2.ed.gov/about/offices/list/oig/auditreports/a03g0006.pdf>
- USDOE, Office of Inspector General. (2007b, March). RMC Research Corporation's Administration of the Reading First Program Contracts: Final Audit Report. USDOE. Retrieved from <http://www2.ed.gov/about/offices/list/oig/auditreports/a03f0022.pdf>
- Valli, L., Croninger, R. G., & Buese, D. (2012). Studying High-Quality Teaching in a Highly Charged Policy Environment. *Teachers College Record*, 114, 1–33.
- Viadero, D., & Manzo, K. K. (2007). Out-of-Favor Reading Plan Rated Highly. *Education Week*, 26(29), 1–14.
- Westat, & Learning Point Associates. (2008, February). Reading First-Ohio: Year 4 Annual Report. USDOE. Retrieved from <http://www2.ed.gov/programs/readingfirst/performance.html>

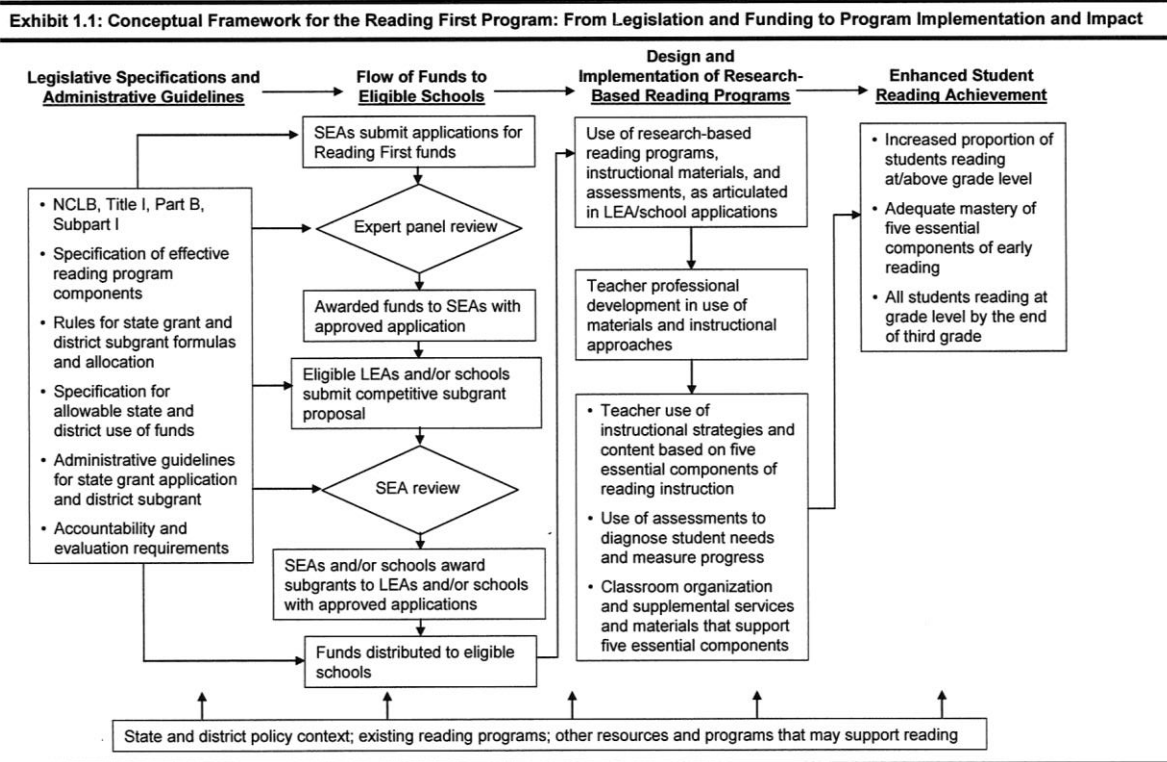
- Willingham, D. T. (2009). *Why don't students like school?: a cognitive scientist answers questions about how the mind works and what it means for the classroom* (1st ed.). San Francisco, CA: Jossey-Bass.
- Wong-Ratcliff, M., Powell, S., & Holland, G. (2010). Effects of the Reading First Program on Acquisition of Early Literacy Skills. *National Forum of Applied Educational Research Journal*, 23(3), 1–14.
- Zhang, Y., & Center on Education Policy, W. (2008). *Some Perspectives from Rural School Districts on the No Child Left Behind Act*. Center on Education Policy.

Appendix A: Summary of Key NCLB Provisions

| Key Provisions of the No Child Left Behind Act | |
|---|--|
| State assessments | States must implement annual state assessments in reading and mathematics in grades 3-8 and at least once in grades 10-12, and in science at least once in each of three grade spans: 3-5, 6-9, and 10-12. Assessments must be aligned with challenging state content and academic achievement standards. States must provide for participation of all students, including students with disabilities and limited English proficient (LEP) students. States must provide for the assessment of English language proficiency of all LEP students. |
| Adequate yearly progress (AYP) | States must set annual targets that will lead to the goal of all students' reaching proficiency in reading and mathematics by 2013-14. For each measure of school performance, states must include absolute targets that must be met by key subgroups of students (major racial/ethnic groups, low-income students, students with disabilities, and LEP students). To make AYP, schools and districts must meet annual targets for each student subgroup in the school, and must test 95 percent of students in each subgroup. States also must define an "other academic indicator" that schools must meet in addition to proficiency targets on state assessments. |
| Schools identified for improvement | Title I schools and districts that do not make AYP for two consecutive years are identified for improvement and are to receive technical assistance to help them improve. Those that miss AYP for additional years are identified for successive stages of interventions, including corrective action and restructuring (see below). To leave identified-for-improvement status, a school or district must make AYP for two consecutive years. |
| Public school choice | Districts must offer all students in identified Title I schools the option to transfer to a non-identified school, with transportation provided by the district. |
| Supplemental educational services | In Title I schools that miss AYP for a third year, districts also must offer low-income students the option of supplemental educational services from a state-approved provider. |
| Corrective actions | In Title I schools that miss AYP for a fourth year, districts also must implement at least one of the following corrective actions: replace school staff members who are relevant to the failure to make AYP; implement a new curriculum; decrease management authority at the school level; appoint an outside expert to advise the school; extend the school day or year; or restructure the internal organization of the school. |
| Restructuring | In Title I schools that miss AYP for a fifth year, districts also must begin planning to implement at least one of the following restructuring interventions: reopen the school as a charter school; replace all or most of the school staff; contract with a private entity to manage the school; turn over operation of the school to the state; or adopt some other major restructuring of the school's governance. Districts must spend a year planning for restructuring and implement the school restructuring plan the following year (if the school misses AYP again for a sixth year). |
| Highly qualified teachers | All teachers of core academic subjects must be "highly qualified" as defined by NCLB and the state. To be highly qualified, teachers must have a bachelor's degree, full state certification, and demonstrated competence in each core academic subject that they teach. Subject-matter competence may be demonstrated by passing a rigorous state test, completing a college major or coursework equivalent, or (for veteran teachers) meeting standards established by the state under a "high, objective uniform state standard of evaluation" (HOUSSE). |
| Use of research based practices | Schools must use effective methods and instructional strategies that are based on scientifically-based research. |

Source: (Robelen, 2002)

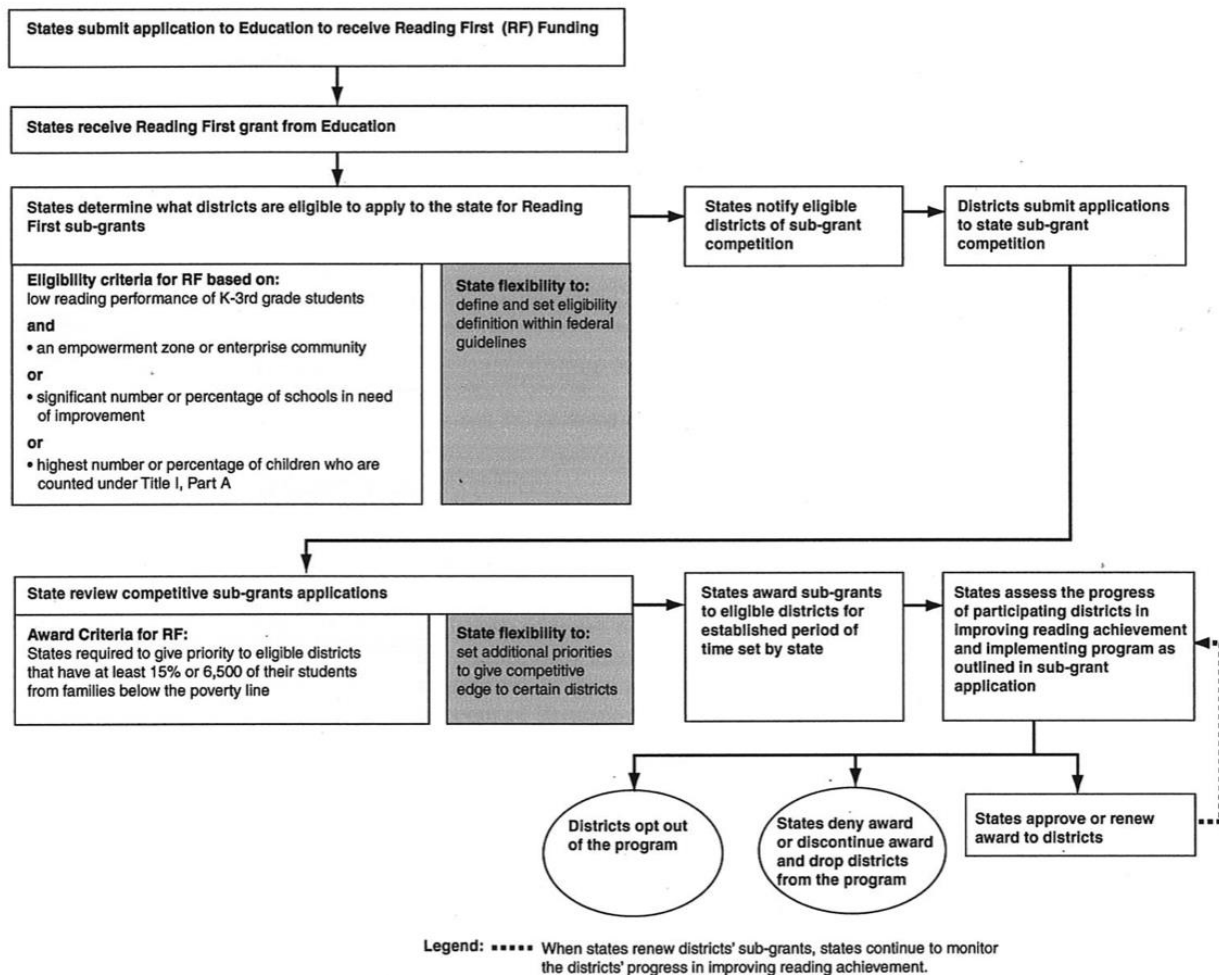
Appendix B: RF Flowchart, Policy to Practice



Source: B. C. Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences (ED), 2008b, p.

Appendix C: Flowchart of State Sub-grants to Districts

Figure 1: The Process for Awarding State Sub-grants to School Districts



Source: GAO analysis based on U.S. Department of Education guidance.

Source: Ashby & General Accounting Office, 2007, p. 11

Appendix D:

Table 1: Reading Program Publishers Most Frequently Represented on States' Approved Lists

| Publisher name | Number of states |
|--|-------------------------|
| Houghton Mifflin ^a | 21 |
| McGraw-Hill-Education ^b | 21 |
| Harcourt ^c | 19 |
| Pearson ^d | 15 |
| Success for All Foundation "Success for All" | 11 |

Source: GAO analysis.

^aBecause some states listed publisher (parent company) names instead of reading program names, our analysis will only reflect overall totals based on publisher names. "The Nation's Choice," a program published by Houghton Mifflin, would be included in this category.

^bMcGraw-Hill Education includes SRA-"Open Court," SRA-"Reading Mastery," and "MacMillan McGraw-Hill Reading."

^cHarcourt, Inc. publishing includes "Trophies."

^dPearson Scott Foresman includes "Scott Foresman Reading," "Longman ESL," and "Scott Foresman Reading Street."

Source: Ashby & General Accounting Office, 2007, p. 18